
Release 3.1a John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

(TM)

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Nov 2 16:31:05 1998; MasPar time 19.22 Seconds
Tabular output not generated. 691.579 Million cell updates/sec

Title: >US-08-906-365-2
Description: (1-530) from US08906365.pep
Perfect Score: 3904
Sequence: 1 MDIKNSPSSINSPPSYNCSQ.....ECSFAPDSKSKGSGSONPQSO 530

Scoring table: PAM 150
Gap 11

Searched: 69111 segs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 50.804; Variance 83.063; scale 0.612

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1698	43.5	589	1	ESTR_CHICK ESTROGEN RECEPTOR (ER)	0.00e+00
2	1693	43.4	595	1	ESTR_HUMAN ESTROGEN RECEPTOR (ER)	0.00e+00
3	1669	42.8	599	1	ESTR_MOUSE ESTROGEN RECEPTOR (ER)	0.00e+00
4	1663	42.6	600	1	ESTR_RAT ESTROGEN RECEPTOR (ER)	0.00e+00
5	1654	42.4	595	1	ESTR_PIG ESTROGEN RECEPTOR (ER)	0.00e+00
6	1603	41.1	583	1	ESTR_OREAU ESTROGEN RECEPTOR (ER)	0.00e+00
7	1591	40.8	575	1	ESTR_ORYLA ESTROGEN RECEPTOR (ER)	0.00e+00
8	1516	38.8	574	1	ESTR_SALIR ESTROGEN RECEPTOR (ER)	0.00e+00
9	1503	38.5	535	1	ESTR_SALISA ESTROGEN RECEPTOR (ER)	0.00e+00
10	807	20.7	433	1	ERRI_HUMAN STEROID HORMONE RECEPTOR (ER)	9.34e-162
11	800	20.5	521	1	ERRI_HUMAN STEROID HORMONE RECEPTOR (ER)	5.20e-160
12	679	17.4	377	1	PRGR_SHEEP PROGESTERONE RECEPTOR	5.25e-130
13	681	17.4	923	1	PRGR_MOUSE PROGESTERONE RECEPTOR	1.69e-130
14	666	17.1	933	1	PRGR_HUMAN PROGESTERONE RECEPTOR	8.35e-127
15	662	17.0	930	1	PRGR_RABIT PROGESTERONE RECEPTOR	8.05e-126
16	615	15.8	776	1	GCR_TUPGB GLUCOCORTICOID RECEPTOR	2.70e-114
17	614	15.7	777	1	GCR_HUMAN GLUCOCORTICOID RECEPTOR	4.74e-114
18	608	15.6	776	1	GCR_XENLA GLUCOCORTICOID RECEPTOR	1.39e-112
19	607	15.5	777	1	GCR_AOTNA GLUCOCORTICOID RECEPTOR	2.43e-112
20	607	15.5	777	1	GCR_SAGOE GLUCOCORTICOID RECEPTOR	2.43e-112
21	587	15.3	786	1	PRGR_CHICK PROGESTERONE RECEPTOR	6.69e-110
22	586	15.0	783	1	GCR_MOUSE GLUCOCORTICOID RECEPTOR	3.20e-107
23	579	14.8	133	1	ESTR_BOVIN ESTROGEN RECEPTOR (ER)	1.62e-105

24	577	14.8	771	1	GCR_CANVO GLUCOCORTICOID RECEPTOR	4.95e-105
25	577	14.8	795	1	GCR_RAT GLUCOCORTICOID RECEPTOR	4.95e-105
26	571	14.6	899	1	ANDR_MOUSE ANDROGEN RECEPTOR	1.42e-103
27	570	14.6	902	1	ANDR_RAT ANDROGEN RECEPTOR	2.48e-103
28	568	14.5	709	1	ANDR_RABIT ANDROGEN RECEPTOR (FRA)	7.59e-103
29	567	14.5	919	1	ANDR_HUMAN ANDROGEN RECEPTOR	1.33e-102
30	562	14.4	981	1	MCR_RAT MINERALOCORTICOID RECEPTOR	2.17e-101
31	544	13.9	758	1	GCR_ONCMY MINERALOCORTICOID RECEPTOR	4.94e-97
32	533	13.7	533	1	RRX_HUMAN RETINOIC ACID RECEPTOR	2.24e-94
33	534	13.7	977	1	MCR_TUPGB MINERALOCORTICOID RECEPTOR	1.28e-94
34	529	13.6	458	1	RRX_MOUSE RETINOIC ACID RECEPTOR	2.06e-93
35	528	13.6	520	1	RRX_MOUSE RETINOIC ACID RECEPTOR	2.06e-93
36	523	13.4	984	1	MCR_HUMAN MINERALOCORTICOID RECEPTOR	5.75e-92
37	517	13.2	111	1	ESTR_SHEEP ESTROGEN RECEPTOR (ER)	1.60e-90
38	514	13.2	467	1	RRX_MOUSE RETINOIC ACID RECEPTOR	8.40e-90
39	514	13.2	467	1	RRX_RAT RETINOIC ACID RECEPTOR	8.40e-90
40	513	13.1	462	1	RRX_HUMAN RETINOIC ACID RECEPTOR	1.46e-88
41	507	13.0	488	1	RRX_XENLA RETINOIC ACID RECEPTOR	4.03e-88
42	502	12.9	597	1	NGFI_RAT NERVE GROWTH FACTOR RECEPTOR	6.37e-87
43	503	12.9	598	1	NGFI_CANFA ORPHAN NUCLEAR RECEPTOR	3.67e-87
44	501	12.9	612	1	MCR_XENLA MINERALOCORTICOID RECEPTOR	3.67e-87
45	501	12.8	598	1	NGFI_HUMAN EARLY RESPONSE PROTEIN	1.11e-86

ALIGNMENTS

RESULT	ID	Query Match	Standard	PRT	Score	Length
1	ESTR_CHICK	43.5%	589 AA			
AC	P06212					
DT	01-JAN-1988 (REL. 06, CREATED)					
DT	01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)					
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)					
DE	ESTROGEN RECEPTOR (ER) (ESTRADIOL RECEPTOR).					
GN	ESR1 OR ESR.					
OS	GALLUS GALIUS (CHICKEN).					
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;					
OC	GALIIFORMES.					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE: 86247578.					
RA	KRUST A., GREEN S., ARGOS P., KUMAR V., WALTER P., BORNERT J.-M.,					
RL	CHAMON P.,					
EMBO	J. 5:891-897 (1986).					
CC	- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.					
CC	- SUBUNIT: HOMODIMER.					
CC	- SUBCELLULAR LOCATION: NUCLEAR.					
CC	- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN, A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.					
CC	- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF TRANSCRIPTIONAL START SITES.					
CC	- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC NUCLEAR HORMONE RECEPTORS.					
DR	EMBL: X03805; G63380; -.					
DR	PIR: S07192; S07192.					
DR	HSSP: P06536; IGDC.					
DR	TRANSFAC: T00264; -.					
DR	PROSITE: PS00031; NUCLEAR RECEPTOR; 1.					
KW	RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN; ZINC-FINGER; STEROID-BINDING.					
FT	DOMAIN 1 178 MODULATING.					
FT	DNA_BIND 179 244 C4-TYPE ZINC FINGERS (TWO).					
FT	ZN_FING 179 199 C4-TYPE.					
FT	ZN_FING 215 239 C4-TYPE.					
FT	DOMAIN 245 304 HINGE.					
FT	DOMAIN 305 589 STEROID-BINDING.					
SEQ	SEQUENCE 589 AA; 66746 MW; 85FF1F4A CRC32;					

Query Match

43.5%; Score 1698; DB 1; Length 589;

ID	ESTR.MOUSE	STANDARD	PRT	599 AA
RESULT	3			
AC	PI9785:			
DT	01-FEB-1991 (REL. 17, CREATED)			
DT	01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	ESTROGEN RECEPTOR (ER) (ESTRADIOL RECEPTOR).			
GN	ESR1 OR ESR OR ESTR OR ESTRA.			
OS	MUS MUSCULUS (MOUSE).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
CC	EUTHERIA; RODENTIA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=UTERUS;			
RX	MEDLINE: 91042558.			
RA	WHITE R., LEES J.A., NEEDHAM M., HAM J., PARKER M.;			
RL	MOL. ENDOCRINOL. 1:735-744(1987).			
CC	-1- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.			
CC	-1- SUBUNIT: HOMODIMER.			
CC	-1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN, A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.			
CC	-1- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING COMPLETELY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF TRANSCRIPTIONAL START SITES.			
CC	-1- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC NUCLEAR HORMONE RECEPTORS.			
CC	EMBL; M38651; G193180; -.			
DR	PIR; A40061; A40061.			
DR	HSSP; P06536; 1GDC.			
DR	TRANSFAC; T00259; -.			
DR	MGI: 109348; ESTRA.			
DR	PROSITE; PS00031; NUCLEAR-RECEPTOR; 1.			
KW	RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;			
KW	ZINC-FINGER; STEROID-BINDING.			
KW	DOMAIN	1	188	MODULATING.
FT	DNA_BIND	189	254	C4-TYPE ZINC FINGERS (TWO).
FT	ZN_FING	189	209	C4-TYPE.
FT	ZN_FING	225	249	C4-TYPE.
FT	DOMAIN	255	599	STEROID-BINDING.
FT	DOMAIN	64	72	POLY-ALA.
SO	SEQUENCE	599 AA;	66955 MW;	6BD9EADB CRC32;
Query Match				
Best Local Similarity 42.8%; Score 1669; DB 1; Length 599;				
Matches 246; Conservative 106; Mismatches 104; Indels 20; Gaps 14				
Db	108	SPSPMLLHPPOLSFPLHHPGQOVVYILENEPSAY-AVRDTGPAPYRSNSDNRRONGR	166	
Oy	73	TTSPVLTPTGHLSP-LVYHROLSHLYAEPOKSPWCBAEASLEHTLPYNNRETLKRVYS-	130	
Db	167	ERLSSNNKKGMMIMESAETRYCAVONCYAGSYHYGWSCGCAAFPRRSIQGNIDYMP	226	
Oy	131	NR-CASPYTG- --GSKDAHFCAVCSYASGYHGWSCGCAAFPRRSIQGNIDYICP	186	
Db	227	ATNCTIDKNNRKSQACRLKRCYEVGMKIGIRKDRGRGMLHKHRODLEGRNEMGA	286	
Oy	187	ATNCTIDKNNRKSQACRLKRCYEVGMKIGIRKDRGRGMLHKHRODLEGRNEMGA	286	
Db	287	SGDMAANLWSPLYIKTKKNSPALSLTADQWYSALEADPEPMIYSEYDPSRPPSEASM	346	
Oy	243	GKAKSGG- -HAPRV-RLILD- -ALS- -PQVLYLTLEAPRPVILS- -RPSAFTTASM	294	

ID	RESULT	4	STANDARD	PRT	600 AA.
AC	ESPR	RAT			
AD	P06211				
DT	01-JAN-1988	(REL. 06, CREATED)			
DT	01-JAN-1988	(REL. 06, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997	(REL. 35, LAST ANNOTATION UPDATE)			
DE	ESTROGEN RECEPTOR (ER) (ESTRADIOL RECEPTOR).				
GN	ESR1 OR ESR OR ESTR.				
OS	RATTUS NORVEGICUS (RAT).				
OC	EDUAROTIA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; RODENTIA.				
NP	[1]				
RP	SEQUENCE FROM N.A.				
RC	SPRAIN-GLIMSTAR;				
RA	MORAMATSU M.;				
RL	SUBMITTED (MAR-1987) TO EMBL/GENBANK/DBJ DATA BANKS.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 871724780.				
RA	KOIKE S., SAKAI M.;				
RL	NUCLEIC ACIDS RES. 15:2499-2513(1987).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	SPRAIN-SPRAGUE-DAWLEY; TISSUE=UTERUS;				
RA	MAGGI A.M.A.;				
RL	SUBMITTED (JUN-1991) TO EMBL/GENBANK/DBJ DATA BANKS.				
CC	-1- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN				
CC	THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR				
CC	PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.				
CC	-1- SUBUNIT: HOMODIMER.				
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.				
CC	-1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,				
CC	A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.				
CC	-1- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT				
CC	TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS: HORMONE BINDING				
CC	GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR				
CC	COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF				
CC	TRANSCRIPTIONAL START SITES.				
CC	-1- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC				
CC	NUCLEAR HORMONE RECEPTORS.				
CC	EMBL; Y00102; G56111; -				
DR	EMBL; X61098; G56121; -				
DR	PIR; S07379; S07379.				
DR	HSSP; P06536; IGDC.				
DR	TRANSFAC; T00258; -				
DR	PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.				
KW	RECEPTOR, TRANSCRIPTION REGULATION, DNA-BINDING, NUCLEAR PROTEIN;				
Y	ZINC-FINGER, STEROID-BINDING.				
FT	DOMAIN	1	189		
FT	DOMAIN_BIND	190	235		
FT	ZN_FING	190	210		
FT	ZN_FING	226	250		
FT	DOMAIN	256	600		
FT	DOMAIN	64	71		
FT	CONFLICT	488	488		
Y	SEQUENCE	600 AA;	67030 MW;	4CCG59B2	CRG32;


```

AC PL16058: 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE ESTROGEN RECEPTOR (ER) (ESTRADIOL RECEPTOR).
GN ESR.
OS SALMO IRIDEUS (RAINBOW TROUT).
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: PISCES: GNATHOSTOMATA:
OC OSTEICHTHYES: ACTINOPTERYGII: SALMONIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91006824.
RA PARDEL F., LE GAC F., LE GAC F., VALOTAIRE Y.,
RL MOL. CELL. ENDOCRINOL. 71:195-204(1990).
RN [2]
RP SEQUENCE OF 150-574 FROM N.A.
RX TISSUE=LIVER.
RA MEDLINE: 89127284.
RA PARDEL F., LE GUELLEC C., VAILLANT C., LE ROUX M.G., VALOTAIRE Y.,
RL MOL. ENDOCRINOL. 3:44-51(1989).
CC -1- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -1- IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT
CC TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS; HORMONE BINDING
CC GREATLY INCREASES RECEPTOR AFFINITY. THE HORMONE-RECEPTOR
CC COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA SEQUENCES UPSTREAM OF
CC TRANSCRIPTIONAL START SITES.
CC -1- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC NUCLEAR HORMONE RECEPTORS.
DR EMBL: M31559; G213784; -.
DR PIR: A37197; A37197.
DR PIR: A40070; A40070.
DR HSSP: P06536; IGDC.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KV ZINC-FINGER; STEROID-BINDING.
FT DOMAIN 1 146 MODULATING (POTENTIAL).
FT DNA_BIND 147 212 C4-TYPE ZINC FINGERS (TWO).
FT ZN_FING 147 167 C4-TYPE.
FT ZN_FING 183 207 C4-TYPE.
FT DOMAIN 213 574 STEROID-BINDING (POTENTIAL).
SQ SEQUENCE 574 AA; 62872 MW; B391ADC2 CRC32;

Query Match 38.8%; Score 1516; DB 1; Length 574;
Best Local Similarity 56.3%; Pred. No. 0.00e+00;
Matches 225; Conservatave 78; Mismatches 75; Indels 22; Gaps 17;

Db 143 ETRYCAVCSDFASGHHYVWVSCSGCAFFKRSIGSHNDYMCPATNOCTMDRRRSQAC 202
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 145 DAHFCAVCSDAVSGHHYVWVSCSGCAFFKRSIGSHNDYICPATNOCTIDKNRRKSQAC 204
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 203 RLRLKCYGVAMVKGGLRKDRGG-RVLRKDKRYCGPAGDREKFTYTWSTGQRPDGGNSS 261
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 205 RLRLKCYGVAMVKGGRSRRRCRGRLVRQ-R--SA-D-EQLHCA-GRAKRS--GGH--AP 253
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 262 SLNCGGGMGRGRITMPPEQVLFLOGQTP-ALCSQKAYARPYTEVTMMTLTSMADKELY 320
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 254 RVREL--LIDA-IS-PEQVLTLLAEPPHVLISHP-A-PFTASMMMSITKLADRELY 307
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 321 HHIAAKVYGPQVQSLSDVOYLLSSWLEVMIGLIRSHCGKLIIFADLLDSESG 380
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 308 HHISIAKKIPGVVELSLDVOYLLSSWLEVMIGLIRSHCGKLIIFADLLDSESG 367
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 381 DCEVMAEITDMLATVSRFGMLKLPPEFVCLAKIILNLNGAFSCSNYSVSLHNSAV 440
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy 366 KVEBIDLEIFDMLATTSRFRELKQKHNEYLCYKAMILLNSMYPVLTATQDA-DSSRKL 426
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 441 ESMLDNITDALIHHTSHSGASVQOOPRRQADLLLLSHIRMSNKGMEHLYSTCKNRKP 500

```

Qy	427	487	501	540	548	552	556	560	564	568	572	576	580	584	588	592	596	600	604	608	612	616	620	624	628	632	636	640	644	648	652	656	660	664	668	672	676	680	684	688	692	696	700	704	708	712	716	720	724	728	732	736	740	744	748	752	756	760	764	768	772	776	780	784	788	792	796	800	804	808	812	816	820	824	828	832	836	840	844	848	852	856	860	864	868	872	876	880	884	888	892	896	900	904	908	912	916	920	924	928	932	936	940	944	948	952	956	960	964	968	972	976	980	984	988	992	996	1000
Y	427	487	501	540	548	552	556	560	564	568	572	576	580	584	588	592	596	600	604	608	612	616	620	624	628	632	636	640	644	648	652	656	660	664	668	672	676	680	684	688	692	696	700	704	708	712	716	720	724	728	732	736	740	744	748	752	756	760	764	768	772	776	780	784	788	792	796	800	804	808	812	816	820	824	828	832	836	840	844	848	852	856	860	864	868	872	876	880	884	888	892	896	900	904	908	912	916	920	924	928	932	936	940	944	948	952	956	960	964	968	972	976	980	984	988	992	996	1000
Y	427	487	501	540	548	552	556	560	564	568	572	576	580	584	588	592	596	600	604	608	612	616	620	624	628	632	636	640	644	648	652	656	660	664	668	672	676	680	684	688	692	696	700	704	708	712	716	720	724	728	732	736	740	744	748	752	756	760	764	768	772	776	780	784	788	792	796	800	804	808	812	816	820	824	828	832	836	840	844	848	852	856	860	864	868	872	876	880	884	888	892	896	900	904	908	912	916	920	924	928	932	936	940	944	948	952	956	960	964	968	972	976	980	984	988	992	996	1000
Y	427	487	501	540	548	552	556	560	564	568	572	576	580	584	588	592	596	600	604	608	612	616	620	624	628	632	636	640	644	648	652	656	660	664	668	672	676	680	684	688	692	696	700	704	708	712	716	720	724	728	732	736	740	744	748	752	756	760	764	768	772	776	780	784	788	792	796	800	804	808	812	816	820	824	828	832	836	840	844	848	852	856	860	864	868	872	876	880	884	888	892	896	900	904	908	912	916	920	924	928	932	936	940	944	948	952	956	960	964	968	972	976	980	984	988	992	996	1000
Y	427	487	501	540	548	552	556	560	564	568	572	576	580	584	588	592	596	600	604	608	612	616	620	624	628	632	636	640	644	648																																																																																								

CC -I- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN


```

CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -1- ALTERNATIVE PRODUCTS: THERE ARE TWO FORMS OF THE PROGESTERONE
CC RECEPTOR: FORM A (THAT STARTS WITH MET-1) AND FORM B (THAT STARTS
CC WITH MET-165).
CC -1- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC NUCLEAR HORMONE RECEPTORS.
DR EMBL: X51730; G35652; -.
DR EMBL: M15716; G189935; -.
DR PIR: A03245; ORHDP.
DR PIR: S09971; S09971.
DR HSP: P06536; IGDC.
DR TRANSFAC: T00696; -.
DR MIM: 264080; -.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR_1.
DR RECEPTOR: TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW ZINC-FINGER; STEROID-BINDING; PHOSPHORYLATION; ALTERNATIVE SPLICING.
FT DOMAIN 1 566
FT DNA_BIND 567 632
FT ZN_FING 567 587
FT ZN_FING 603 627
FT DOMAIN 681 933
FT MOD_RES 183 187
FT MOD_RES 41 41
FT MOD_RES 227 227
FT MOD_RES 232 232
FT MOD_RES 552 552
FT MOD_RES 793 793
FT VARSLIC 1 164
FT CONFLICT 226 226
FT CONFLICT 256 256
FT CONFLICT 344 344
FT CONFLICT 660 660
SQ SEQUENCE 933 AA; 99042 MW; 0D1124EB CRC32;

Query Match 17.1%; Score 666; DB 1; Length 933;
Best Local Similarity 33.4%; Pred. No. 8.35e-127;
Matches 110; Conservative 89; Mismatches 113; Indels 17; Gaps 17;

DB 566 ICILICDEASGCHYGLTGSCVFFRKAMEGHNILCAGRNDCIYDKIRKNCPCACRLR 625
QY 148 FCVAVCSDYASGYHYGWSCGCAFFKRSIQGHNDYICATNCTIDKNRKCQACRLR 207
DB 626 KCCQAGNVL-GGRKFK-KENKRVVRALDAVALPQVIGIPNESQ-RITSPSOEIQ-LIP 662
QY 208 KCYEVMVWCGSRRCRGYRLVRRORSAD-QL-HCAGAKRSGGHAPRVRELLDALSP 265
DB 684 IP-PLINLMSIEPVIYAGHDNTPDSSSLTSLNQLGEROLLVYKWSKSLPGRNL 742
QY 264 SPQVLTLLEAPPHYLISRSA-PFTESMMMSLTKLADKELVHMSAKKIPGVEL 322
DB 743 HIDDITLLOYSMWSLWFGWRSYKRVSGQMLYFADLLINEQRMK-ESSFYSICLTLM 801
QY 323 SLFDQVRLLESCEMVEYLMGLMWSIDH-PGKLI-FAPDLVLDREDEKCEVGIETFDML 380
DB 802 WQDPQEFKQVSOEFLCMVLLNLTIPLEGRSOTO-FEEMRS-SYI-RELIKAIG- 857
QY 381 LATTSFRELKQHKREYLCVAMILLNS-SMYPVATATDADSSRLAHLLNAVTDALVW 439
DB 858 LROGVYSSQRFYQTLTKLDNLHDLVKQ 886
QY 440 VIAXSGISSQOSMRLANLL-MLSHYRH 467

```

```

RESULT 15
ID PRGR_RABIT STANDARD; PRT; 930 AA.
AC P06186;
DT 01-JAN-1988 (REL. 06, CREATED)
DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE PROGESTERONE RECEPTOR (PR).
GN PGR.

```

```

OS ORYCTOLAGUS CUNICULUS (RABBIT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; LAGOMORPHA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 87067449.
RA LOOSELY F., ATGER M., MISRAHI M., GUICHON-MANTEL A., MERIEL C.,
RA LOOSELY F., BENAROUS R., MILGROM E.;
RL PROC. NATL. ACAD. SCI. U.S.A. 83:9045-9049(1986).
CC -1- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -1- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC
CC NUCLEAR HORMONE RECEPTORS.
DR EMBL: M14547; G165632; -.
DR PIR: A25923; A25923.
DR HSP: P06536; IGDC.
DR TRANSFAC: T00697; -.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR_1.
DR RECEPTOR: TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW ZINC-FINGER; STEROID-BINDING.
FT DOMAIN 1 565
FT DNA_BIND 566 633
FT ZN_FING 566 588
FT ZN_FING 604 628
FT DOMAIN 678 930
FT DOMAIN 184 188
SQ SEQUENCE 930 AA; 98666 MW; EE133E86 CRC32;

Query Match 17.0%; Score 662; DB 1; Length 930;
Best Local Similarity 32.5%; Pred. No. 8.05e-126;
Matches 106; Conservative 91; Mismatches 114; Indels 15; Gaps 15;

DB 567 ICILICDEASGCHYGLTGSCVFFRKAMEGHNILCAGRNDCIYDKIRKNCPCACRLR 626
QY 148 FCVAVCSDYASGYHYGWSCGCAFFKRSIQGHNDYICATNCTIDKNRKCQACRLR 207
DB 627 KCCQAGNVL-GGRKFK-KENKRVVRALDAVALPQVIGIPNESQ-RITSPSOEIQ-LIP 662
QY 208 KCYEVMVWCGSRRCRGYRLVRRORSAD-QL-HCAGAKRSGGHAPRVRELLDALSP 265
DB 683 -PLINLMSIEPVIYAGHDNTPDSSSLTSLNQLGEROLLVYKWSKSLPGRNLH 741
QY 266 EQVLTLLEAPPHYLISRSA-PFTESMMMSLTKLADKELVHMSAKKIPGVELSL 324
DB 742 DDITLLOYSMWSLWFGWRSYKRVSGQMLYFADLLINEQRMK-ESSFYSICLTLMQ 800
QY 325 FQVYRLLESCEMVEYLMGLMWSIDH-PGKLI-FAPDLVLDREDEKCEVGIETFDMLA 382
DB 801 IPOEFYKQVSOEFLCMVLLNLT-I-PLERLSQSOFEEMRSYIRELIKAIG-LRQ 857
QY 383 TTSFRELKQHKREYLCVAMILLNS-SMYPVATATDADSSRLAHLLNAVTDALVWYA 442
DB 858 KGVYSSQRFYQTLTKLDNLHDLVKQ 883
QY 443 KSGISSQOSMRLANLL-MLSHYRH 467

```

Search completed: Mon Nov 2 16:32:02 1998
Job time : 57 secs.

THIS PAGE BLANK IS 19

M O S E R H
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msrch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Nov 3 04:56:49 1998; MasPar time 74.31 Seconds

Tabular output not generated. 1058.292 Million cell updates/sec

Title: >US-08-906-365-1
Description: (1-1686) from US08906365.seq
Perfect Score: 1686
N.A. Sequence: 1 cagccattactgtccac.....agaccacagctcagtgta 1686
Comp: gtcgtatatagaacgggtg.....tcttggtgctcagtgctact

Scoring table: TABLE default
Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 88822 seqs, 23323279 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-issued
1:5_COMB 2:PCT9_COMB 3:backfile1

Statistics: Mean 9.055; Variance 4.480; scale 2.021

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	197	11.7	4963	1	US-08-076-Sequence 16, Applicat	2.47e-138
2	197	11.7	4963	1	US-08-260-Sequence 9, Applicatio	2.47e-138
3	71	4.2	3460	1	US-08-312-Sequence 1, Applicatio	2.34e-35
4	63	3.7	6244	1	US-08-076-Sequence 15, Applicati	3.17e-29
5	63	3.7	6244	1	US-08-260-Sequence 8, Applicati	3.17e-29
6	62	3.7	7218	1	US-08-232-Sequence 14, Applicati	1.82e-28
7	61	3.6	7218	1	US-08-232-Sequence 14, Applicati	1.04e-27
8	55	3.3	2989	3	5223606-1 Patent No. 5223606.	3.26e-23
9	53	3.1	3014	1	US-08-759-Sequence 1, Applicatio	9.86e-22
10	53	3.1	3014	1	US-08-629-Sequence 1, Applicatio	9.86e-22
11	51	3.0	1576	3	5260432-1 Patent No. 5260432.	2.91e-20
12	47	2.8	387	1	US-08-087-Sequence 1, Applicatio	2.34e-17
13	47	2.8	387	1	US-08-561-Sequence 1, Applicatio	2.34e-17
14	42	2.5	2928	2	PCT-US92-0 Patent No. 5171671.	8.39e-14
15	42	2.5	2940	3	5171671-1 Patent No. 5171671.	8.39e-14
16	42	2.5	3036	2	PCT-US91-0 Sequence 1, Applicatio	4.95e-11
17	38	2.3	1662	2	US-08-238-Sequence 5, Applicatio	5.22e-09
18	35	2.1	215	1	US-08-238-Sequence 5, Applicatio	1.10e-07
19	33	2.0	215	1	US-08-338-Sequence 1, Applicatio	1.10e-07
20	33	2.0	1407	1	US-08-459-Sequence 1, Applicatio	1.10e-07

21	33	2.0	1952	1	US-08-463-Sequence 1, Applicatio	1.10e-07
22	33	2.0	1952	1	US-08-333-Sequence 3, Applicatio	1.10e-07
23	30	1.8	2095	2	PCT-US91-0 Sequence 1, Applicatio	9.42e-06
24	30	1.8	2130	1	US-07-952-Sequence 1, Applicatio	9.42e-06
25	29	1.7	1866	2	PCT-US91-0 Sequence 1, Applicatio	4.01e-05
26	28	1.7	2204	1	US-07-952-Sequence 3, Applicatio	1.67e-04
27	28	1.7	2219	2	PCT-US91-0 Sequence 7, Applicatio	1.67e-04
28	28	1.7	2468	1	US-08-333-Sequence 11, Applicati	1.67e-04
29	28	1.7	2468	1	US-08-463-Sequence 1, Applicati	1.67e-04
30	25	1.5	74	2	PCT-US95-1 Sequence 94, Applicati	1.06e-02
31	25	1.5	75	2	PCT-US95-1 Sequence 99, Applicati	1.06e-02
32	25	1.5	81	2	PCT-US95-1 Sequence 92, Applicati	1.06e-02
33	25	1.5	82	2	PCT-US95-1 Sequence 97, Applicati	1.06e-02
34	25	1.5	82	2	PCT-US95-1 Sequence 97, Applicati	1.06e-02
35	26	1.5	242	1	US-08-273-Sequence 1, Applicatio	2.73e-03
36	26	1.5	1659	2	PCT-US94-0 Sequence 1, Applicatio	2.73e-03
37	26	1.5	1659	1	US-08-463-Sequence 7, Applicatio	2.73e-03
38	26	1.5	1659	1	US-08-333-Sequence 7, Applicatio	2.73e-03
39	26	1.5	2304	1	US-08-464-Sequence 1, Applicatio	2.73e-03
40	26	1.5	2304	1	US-08-464-Sequence 1, Applicatio	2.73e-03
41	24	1.4	69	1	US-08-471-Sequence 142, Applicat	4.05e-02
42	24	1.4	74	2	PCT-US95-1 Sequence 100, Applicat	4.05e-02
43	24	1.4	81	2	PCT-US95-1 Sequence 98, Applicati	4.05e-02
44	24	1.4	81	2	PCT-US95-1 Sequence 92, Applicati	4.05e-02
45	24	1.4	2009	1	US-08-463-Sequence 9, Applicatio	4.05e-02

ALIGNMENTS

RESULT 1
ID US-08-076-726-16 STANDARD; DNA; UNC; 4963 BP.
AC xxxxxx

DE Sequence 16, Application US/08076726
CC Sequence 16, Application US/08076726
CC Patent No. 5464758

GENERAL INFORMATION:

CC APPLICANT: Gossen, Manfred
CC APPLICANT: Bujard, Hermann
CC TITLE OF INVENTION: Tight Control of Gene Expression in
CC NUMBER OF SEQUENCES: 16
CC Eucaryotic Cells by Tetracycline-responsive Promoters
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Sterne, Kessler, Goldstein & Fox
CC STREET: 1100 New York Avenue, N.W.
CC CITY: Washington
CC STATE: District of Columbia
CC COUNTRY: United States of America
CC ZIP: 20005-3934

COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC SOFTWARE: PC-DOS/MS-DOS
CC OPERATING SYSTEM: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/076,726
CC FILING DATE: 14-JUN-1993

CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:
CC NAME: Esmond, Robert W.
CC REGISTRATION NUMBER: 32,893
CC REFERENCE/DOCKET NUMBER: 0942.2490001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202)371-2600
CC TELEFAX: (202)371-2540

INFORMATION FOR SEQ ID NO: 16:

CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 4963 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: both
CC TOPOLOGY: both
CC SEQUENCE 4963 BP; 1259 A; 1300 C; 1278 G; 1126 T; 0 OTHER.

Qy	665	gtacatcgtacaaaccgcgcgaagctctccagcctctccgacttclogaagatgttacg	725
Db	2545	AAGTCGCGCATGGT	2557
Qy	725	aagtcgcatgctgt	737

RESULT	5
ID	US-08-260-452-8 STANDARD; DNA; UNC; 6244 BP
AC	xxxxxx

DE Sequence 8, Application US/08260452
CC Sequence 8, Application US/08260452
CC Patent No. 5650298
CC
CC GENERAL INFORMATION:
CC APPLICANT: Gossen, Manfred
CC APPLICANT: Bujard, Hermann
CC APPLICANT: Salfeld, Jochen
CC APPLICANT: Voss, Jeffrey
CC TITLE OF INVENTION: Tight Control of Gene Expression in Eucaryotic
CC TITLE OF INVENTION: Cells by Tetracycline-responsive Promoters
CC NUMBER OF SEQUENCES: 10
CC
CC CORRESPONDENCE ADDRESSES: 10

```
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: ASCII text
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/260,452
CC FILING DATE:
```

CC INFORMATION FOR SEQ ID NO: 8:

CC MOLECULE TYPE: DNA (genomic)
CC ORIGINAL SOURCE:
CC ORGANISM: Human cytomegalovirus
CC STRAIN: Towne (hcmv)
CC IMMEDIATE SOURCE:
CC

Query Match	3.78;	Score 63;	DB 1;	Length 6244;
Best Local Similarity	66.3%;	Pred. No. 3.17e-29;		
Matches 128:	Conservative	0;	Mismatches 65;	Indels 0;
				Gaps 0

Db 2365 TCGTGGGAGTGAAGCATCAGGCGTGATTTATGGTGTCCACCTGTGGAGCTGTGAAG 2422
Oy 545 ttctgaagcagatcaagcatcgagataatacaactcagtgagctgctgtgtgaaagatgtaag 604
Db 2425 TCTTCTTTAAAAGGCAATGAGAGGAGCATACTCTTTATGTGCTGGAGAAATGACT 2488
Oy 605 cctttttttaaagaagccttcaagaacataatgattatcttccagctacacacagtt 664

[illegible]

RESULT	6
ID	US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP
AC	xxxxxx

DE Sequence 14, Application US/08232243
CC Sequence 14, Application US/08232463
CC Patent No. 5670367
CC GENERAL INFORMATION:
CC APPLICANT: DORNER, F.
CC APPLICANT: SCHIEFLINGER, F.
CC APPLICANT: FALKNER, F. G.
CC TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
CC NUMBER OF SEQUENCES: 52
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/232,463
CC FILING DATE:

TELECOMMUNICATION INFORMATION

```
CC LENGTH: 7218 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
```

Query Match	3.7%;	Score 62;	DB 1;	Length 7218;
Best Local Similarity	1.4%;	Pred. No. 1.82e-28;		
Matches	5;	Conservative	205;	Mismatches 148;
			Indels	0;
			Gaps	0;

Db 1083 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1144
 Cp 864 ctccgcacactcgggcgcgfcgcgcacactctctgcttgcgcgaacagtgcagctg 805
 Db 1143 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1204
 Cp 804 ctccgcgcacactctctcgtctccgcgaacaggcgtatccacacatctctctccgcggagcc 745

[illegible]

QY	656	caatcagtcgtaacatcgataaaacccggcgcaagagctgcgcagcctgcgcgaattcoga	715
Db	367	AATGTCCTTCAGGCTGAATG	386
QY	716	agtyrtacgaagtgggaatg	735

RESULT	14
ID	PCT-US92-02320A-3 STANDARD; DNA; UNC; 2928 BP
AC	xxxxxx

DE Sequence 3, Application PC/TUS9202330A
DE Sequence 3, Application PC/TUS9202330A
CC GENERAL INFORMATION:
CC APPLICANT: Sloan-Kettering Institute, For Cancer Research
CC TITLE OF INVENTION: METHODS FOR DETECTION AND TREATMENT OF CANCER
CC NUMBER OF SEQUENCES: 6
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: COOPER & DUNHAM
CC STREET: 30 ROCKEFELLER PLAZA
CC CITY: NEW YORK
CC STATE: NEW YORK
CC COUNTRY: U.S.A.
CC ZIP: 10112
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent'n Release #1.24
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US92/02320A
CC FILING DATE:

```

CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: 56 675,084
CC      FILING DATE: 22-MAR-1991
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: 56 673,838
CC      FILING DATE: 22-MAR-1991
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: WHITE, JOHN P
CC      REGISTRATION NUMBER: 28,678
CC      REFERENCE/DOCKET NUMBER: 38694-PCY
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (212) 977-9550
CC      TELEFAX: (212) 644-0525
CC      TELEX: (212) 422523 COOP UT
CC      INFORMATION FOR SEQ ID NO: 3:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 2928 base pairs
CC      TYPE: nucleic acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: cDNA
CC      IMMEDIATE SOURCE:
CC      CLONE: hRAR ALPHA
CC      FEATURE:
CC      NAME/KEY: CDS
CC      LOCATION: 103..1488
CC      OTHER INFORMATION:
SQ      SEQUENCE 2928 BP: 562 A; 1011 C; 789 G; 546 T; 0 OTHER.

```

[illegible]

RESULT 15
 ID 5171671-1 STANDARD; DNA; UNC; 3185 BP.
 AC xxxxxx
 DT 01-JAN-1900
 DE Patent No. 5171671.
 CC Patent No. 5171671
 CC APPLICANT: EVANS, RONALD M.; ONG, ESTELLITA S.; SEGUI,
 CC PRUD'HMAR S.; THOMPSON, CATHERINE C.; UEMSONO, KAZUHIKO
 CC GUGIERE, VINCENT
 CC TITLE OF INVENTION: RETINOIC ACID RECEPTOR COMPOSITION
 CC NUMBER OF SEQUENCES: 2
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/07/546,256
 CC FILING DATE: 06-AUG-1990
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 276,536
 CC FILING DATE: 30-NOV-1988
 CC APPLICATION NUMBER: 128,331
 CC FILING DATE: 02-DEC-1987
 CC SEQ ID NO:1:
 CC LENGTH: 2940
 SQ Sequence 3185 BP: 594 A; 1010 C; 790 G; 546 T; 245 other;

	Query Match	2.5%;	Score 42;	DB 3;	Length 2940;
	Best Local Similarity	60.6%;	Pred. No. 8.39-14;		
	Matches 120;	Conservative	Mismatches 0;	Indels 78;	Gaps 0;
Db	364	TGCTTTGCTGTGCAGACAGACTCCTCAGGCTACCATATGGGGCTACGCGCTGTGAGGC	423		
Qy	538	tgcgctgtgctcgaacgatacgcatacggataactatgagctgtgctgtgaagga	597		
Db	424	TGCAAGGGCTCTTTCGCCGCGCAGATCCAGAAACAATGGTGTCACAGTGTGCACCGGAC	483		
Qy	558	tgtaagagccctttttaaagaagaatccaagaataatgtattatlttcgcagctaca	657		
Db	484	AAGAACTGCATCATCAACACAGGTATACCCGGAAACCGTCCAGTACTGCCGACATGCACAG	543		
Qy	658	aatcagtgtaaatcgtataaanaacccgycgaagagctgcagcgctgcagacttcggaag	717		
Db	544	TGCTTTGAAGTGGGCGATG	561		
Qy	718	tgtaagaagctgggaatg	735		

Search completed: Tue Nov 3 04:58:08 1998
Job time : 79 secs.

 N O T E S

 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch_n n.a. - n.a. database search, using Smith-Waterman algorithm
 Run on: Tue Nov 3 03:32:21 1998; MasPar time 2544.39 Seconds
 1354.241 Million cell updates/sec
 Tabular output not generated.

Title: >US-08-906-365-1
 Description: (1-1686) from US08906365.seq
 Perfect Score: 1686
 N.A. Sequence: 1 cagccattactctgccac.....agaccacagctcagctga 1686
 Comp: gtcggtatataagcaggggtg.....tctgggtgctcagctcact

Scoring table: TABLE default
 Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 552174 seqs, 1021863385 bases x 2

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: emb155

1:em_db 2:em_htg 3:em_hum1 4:em_hum2 5:em_in 6:em_com
 7:em_ov 8:em_ov 9:em_pat 10:em_pl 11:em_ro
 genbank107

12:gb_db 13:gb_htg 14:gb_in 15:gb_com 16:gb_ov 17:gb_pat
 18:gb_ph 19:gb_pl 20:gb_pri 21:gb_pri 22:gb_ro 23:gb_st
 24:gb_sts 25:gb_sy 26:gb_un 27:gb_vt

Statistics: Mean 11.383; Variance 4.965; scale 2.293

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1684	99.9	1740	21	AB006590 Homo sapiens mRNA for	0.00e+00
2	1683	99.8	2011	21	AF051427 Homo sapiens estrogen	0.00e+00
3	1496	88.7	2041	21	AF051428 Homo sapiens estrogen	0.00e+00
4	1495	88.7	2745	21	AF060555 Homo sapiens estrogen	0.00e+00
5	1495	88.7	3593	21	AB006589 Homo sapiens mRNA for	0.00e+00
6	1452	86.1	1460	9	A61586 Sequence 4 from Patent	0.00e+00
7	1450	86.0	1560	20	HSRNERB H.sapiens mRNA for est	0.00e+00
8	1140	67.6	1650	22	RNAJ2602 Rattus norvegicus mRNA	0.00e+00
9	1136	67.4	2555	9	A61583 Rattus norvegicus estr	0.00e+00
10	1036	61.4	1621	22	AF042058 Rattus norvegicus estr	0.00e+00
11	1021	60.6	2152	22	MM081451 Mus musculus estrogen	0.00e+00
12	1021	59.9	1458	9	A61388 Sequence 6 from Patent	0.00e+00
13	1010	59.9	1458	22	MM081451 Mus musculus mRNA for	0.00e+00
14	1010	59.9	1458	22	MM081451 Mus musculus estrogen	0.00e+00
15	781	46.3	1215	21	AF074599 Homo sapiens estrogen	0.00e+00

16	720	42.7	1512	22	RNAJ2603 Rattus norvegicus mRNA	0.00e+00
17	720	42.7	1675	22	AF042059 Rattus norvegicus estr	0.00e+00
18	716	42.5	1512	22	AB012721 Rattus norvegicus mRNA	0.00e+00
19	656	38.9	1504	22	AF042060 Rattus norvegicus estr	0.00e+00
20	370	21.9	1558	22	AF042061 Rattus norvegicus estr	0.00e+00
21	336	19.9	3061	16	AB003356 Bel mRNA for estrogen	3.93e-282
22	277	16.4	332	21	CJERB C.jacchus mRNA for est	1.76e-225
23	221	13.1	306	21	AF074598 Homo sapiens estrogen	3.57e-172
24	217	12.9	2770	16	ORZMR Oryzias sp. mRNA for e	2.17e-168
25	203	12.0	1788	15	SSESTREC S.scrota mRNA for estr	3.57e-155
26	203	12.0	3500	16	AB007453 Chrysophrys major mRNA	3.57e-155
27	201	11.9	1966	15	O.aries mRNA for estro	2.73e-153
28	201	11.9	2085	16	GEER Chicken mRNA for estro	2.73e-153
29	197	11.7	2092	17	IO8538 Sequence 1 from Patent	1.59e-149
30	197	11.7	2092	20	HMEBRCF Human estrogen recepto	1.59e-149
31	197	11.7	2322	17	AA2099 Sequence 1 from Patent	1.59e-149
32	197	11.7	4963	17	I15368 Sequence 16 from patent	1.59e-149
33	197	11.7	4963	17	I56760 Sequence 9 from patent	1.59e-149
34	197	11.7	6450	21	HSER Human mRNA for oestrog	1.59e-149
35	189	11.2	2792	16	ZFIER Taenioygyia guttata es	5.21e-142
36	184	10.9	1880	22	RNESTROR Rat mRNA for estrogen	2.54e-137
37	184	10.9	2090	22	RNESTROR Rat mRNA for estrogen	2.54e-137
38	184	10.9	2090	22	RNESTROR Rat mRNA for estrogen	2.54e-137
39	182	10.8	372	21	AF061055 Homo sapiens estrogen	1.89e-135
40	182	10.8	659	21	AF061054 Homo sapiens estrogen	1.89e-135
41	178	10.6	2078	22	MUSESTRU Mouse estrogen recepto	1.04e-131
42	175	10.4	586	20	HS275126 H.sapiens mRNA for est	6.62e-129
43	172	10.2	1662	16	SSESTREC S.salar mRNA for oest	4.17e-126
44	170	10.1	1278	16	SMOERA Rainbow trout estrogen	3.06e-124
45	164	9.7	828	16	AF013104 Sparus aurata estrogen	1.17e-118

ALIGNMENTS

RESULT	1	LOCUS	1740 bp	RNA	PRI	27-FEB-1998
DEFINITION	AB006590	Homo sapiens mRNA for estrogen receptor beta, complete cds.				
ACCESSION	AB006590					
KEYWORDS	92911151	estrogen receptor beta.				
SOURCE		Homo sapiens CDNA to mRNA.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;				
AUTHORS		1 (sites) Ogawa,S., Inoue,S., Watanabe,T., Hiroi,H., Otsu,A., Hosoi,T.,				
TITLE		Ouchi,Y. and Muramatsu,M.				
JOURNAL		The complete primary structure of human estrogen receptor beta (ER				
MEDLINE		beta) and its heterodimerization with ER alpha in vivo and in vitro				
REFERENCE		Biochem. Biophys. Res. Commun. 243 (1), 122-126 (1998)				
AUTHORS		2 (bases 1 to 1740)				
TITLE		Ogawa,S.				
JOURNAL		Submitted (13-AUG-1997) to the DDBJ/EMBL/Genbank databases. Sumitro				
MEDLINE		Ogawa, Saitama Medical School, Department of 2nd Biochemistry, 38				
REFERENCE		Morohongo, Moriyama, Iruma-gun, Saitama 350-04, Japan				
AUTHORS		(E-mail: suogawa@saitama-med.ac.jp, Tel: 81-492-76-1490,				
TITLE		Fax: 81-492-94-9751)				
FEATURES		Location/Qualifiers				
source		1..1740				
CDS		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
		99..1691				
		/codon_start=1				
		/product="estrogen receptor beta"				
		/db_xref="pid:d1025878"				
		/db_xref="pid:g2911152"				
		/translation="MDIKNSPSSLSPPSSYINCQSILPLEHSGIYIPSSYVDSHNEYP				
		AMTFYSPAYWNSIPSNVNLLEGPGRTQSPNVLPFGHSLPLVHRLSLTAEF				
		OKSPWCFAVSLHTLPLVNRRLTKRKVSNCAFPYGPSKRAHFCVACSDVASGYH				
		YGVASCEGKAFKRSIOGDHNDYICATNOCTIDKRRRSQCOGCRKCYEVMWVCG				
		SRRERCGYLVRQRADQLHCGAKKSGGAPRVRELLDALSPDQLVLTLEAE				

BASE COUNT 425 a 463 c 459 g 393 t
ORIGIN

BASE COUNT	425 a	463 c	459 g	393 t
ORIGIN				

Query Match	99.9%;	Score 1684;	DB 21;	Length 1740;
Best Local Similarity	99.9%;	Pred. No. 0.00e+00;		
Matches 1685; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

D	6	CAGCAGTATATCTGTGCCAGAAATCTTTGGAACTATATATACCTTTTGCTCTTC	65
Q	1	cagccattatactctgcgccacgaaccttttggaataataatgaccttttgccctctc	60
D	66	TGCAGAGTGTTTTCTCAGCTGTATTCTCAAGACATGGATATAAAAAACTCACCATCTAGC	125
Q	61	tgcgaagtggtttcttcagctgtctatctccaagatcgtatataaataccacatctagc	120
D	126	CTTATTTCTCTCTCTCTCTACACTGCAATCATCTTACCCTTGAGACAGCGCTCC	185
Q	121	cttaattctctctctcccttaaacatgcaftcaatccattacccttgcagcagctcc	180
D	186	ATTATACATACCTTCTCTCTATGTGAAGGACCATTAATATCCAGCATGACATCTAT	245
Q	181	atatacaactcttccctcttaqtaagacagccacaatgaataccagccaagaattctat	240
D	246	AGCCCTGCTGTGATGAATTACAGATTCGCCAGCATCTCACTTACTTGGAAAGGTGGGCT	305
Q	241	agccctgtgtgatagaattaaagatctccagcaatgtaacttaacttggaagtggcct	300
D	306	GGTGGGACACACAAAGCCCAATGTGTGTGGCCAACCTGGCACCTTCTCTTTA	365
Q	301	gtctgcgacagacacaagaagcccaatgtgtgtgccaacaccttgcgaaccttctctta	360
D	366	GTGGTCATTCGCCAGTTTATCATCTGTATGCGGAACCTCAAAAAGTCTCCTGTGTA	425
Q	361	gtgtgtccatctgcagattatcatctctlaatgcygaacctcaaaagatccctgtgtga	420
D	426	GCAATATGCTATGAACACACCTTACCTGTAAACAGAGAGACACTGTAAAGAAAGTGTAT	485
Q	421	gcaaatatctgcagaaacacaacctcttcctgtaaacgagaagacactgaaagaagtgta	480
D	486	GGGAACCGTTGGCCAGCCCTGTACTGTGTCAAGTTCAAGAAGATGCTCACTTCTGC	545
Q	481	gggaaccgctggccagccctgttaactgtctcaggttccaagaaggagtgtctactctgc	540
D	546	GCTGTCTGACAGGATTACGCATCGGGAATATACATATGAGTGTGTGTGTAAGATGT	605
Q	541	gtctgtctcagagatctacgcacatcggaatatacctaatactgtgtctgtgaaagat	600
D	606	AAGGCTTTTAAAGAAGATTCAAAGACATATATATATTTTCCAGCTACAAAT	665
Q	601	aaggcctctttaaagaagaacatccaagaataatgatatatttccagctacaat	660
D	666	CAGTGTACATCGATMAAACCGGGCGAAGACTCCAGGCTGCCACTTCGAAATGT	725
Q	661	cagtgtaacatcgataaanaacggcgcaagctgcgaagcgcttcggaatgt	720
D	726	TACGAATGGGAATGCTGAATGTGTGCTCCCGGAGAGAGAGATGTGGGTACCGCTTTGT	785
Q	721	tacgaatgtygaatgtygaatgtgtgtctcccggaagagaatgttgygttaccgcttgt	780
D	786	CGGAGACAGAGAGTGGCGAGACAGTGTGACGTGGCGGCAAGGCGCAGAGAGAAGTGG	845
Q	781	cggagaacagaagaatgycgcgaagagaagctgcactgtgcygcaagcgcaagaagaatgyc	840
D	846	GGCAGAGCCCGGAGTGGGAGCTGTCTGTGAGCCGCTGAGACCCCGAGCAGCTAGT	905
Q	841	ggcagagcgccccgagtgcyggaagctgtgtctgtagagccctgagcccccgagcagctag	900
D	906	CTCACCTCTTGAGAGCTGAGCCGCCCATATGCTGATCAGCGGCCCATGCGACCTTC	965

Qy	901	ctcaacctctctgagaggtctgagccgccccatgctctgatacagccgccccagctgccccttc	960
Db	966	ACCGAGGCTTCATGATGATGATGTCCTCTBACCAATTGGCCGACAGAAGATTGGTATACATG	1025
Qy	961	accgagagccttcacatgatagtaatctctcttccacaagtctggccgacaagagatctgttaacacag	1020
Db	1026	ATCAGCTGGGGCAGAGAGATTTCCCGGCTTTGTGTGAGCTCAGCCTCTTTCACACAAGTGGCG	1085
Qy	1021	atcagctgaggccaagaagaatctcccggtcttgtagctcagctcgtctgttcacaaagtcgag	1080
Db	1086	CTCTTGGAGAGCTTGTGGATGAGAGTGTATATATGAGGGCGCTATGTGGCGCTCATTTGAC	1145
Qy	1081	ctctttagagactctgtgatacgagagtgcttaatactatgtaggtctgtagctgtgcctaaattgac	1140
Db	1146	CACCCGGCAGAGCTCATCTTTGGCTCCGAGATCTTGTCTGTGACACAGGGATGAGGGGAATGC	1205
Qy	1141	cacccgcgaagcctcaactcttctcctcagatcctgtctctgacagggatagaagggaaatgc	1200
Db	1206	GTAGAAGGAATTTGGAATCTTTTACATGCTCCTGGCAACTACTTTCAGAGTTTCAGAG	1265
Qy	1201	gtagaaggaattctggaatactctttagaatctctcctgcaactcctcaaggtcttcagag	1260
Db	1266	TTAAACATCCCAACAAAGAAATATCTGTGTGTAAGGCCATGATCCCGTCATTCACAT	1325
Qy	1261	ttaaacatcccaacaaaagaatactctctgtcacaaggccaatgctcctgcctcaattccagct	1320
Db	1326	ATGTACCTCTGTGTACAGAGCGACCACAGATGCTGACAGCACGCCGGAAGCTGACTCTTG	1385
Qy	1321	atgtacacctctgtcacaagcgaccagagatgctgtacagcagccggaagctgtgctcaacttg	1380
Db	1386	CTGAACGCCGTCGACCATGCTTTGGTTGGGTGATTTGCCAAGACGGCATCTCTCCACG	1445
Qy	1381	ctgaacgcgcgtgacccgatgtcttggtttggtgtgtctgccaagacgagcactctccacag	1440
Db	1446	CAGCATCTCATGCGCCGTCGGCTTAACCTCTGATGCTCTGTGCCACGTCAGCATGCGAGT	1505
Qy	1441	cagcatctcacaagcgcctcgtcctaactctctgactgtcctgtcccaagctcagagctgagatg	1500
Db	1506	AACAAGGCGATGAGACATCTGCTCACAATGATGATGCAAAATGTGTCCAGTGTATGAC	1565
Qy	1501	aacaagggcatggaacatctgtctcaacatgaagatgacaaatgtgtgtccagtgatgac	1560
Db	1566	CTGCTGCTGGAGATGCTGAATGCCACGCTGTGGCGGGTGAAGTCTCTCATACAGGGG	1625
Qy	1561	ctgctgctggagatgactgtcgaatgcccacgctgtctcgcggtgcgaagctctccatccacgggg	1620
Db	1626	TCCGAGTCACAGCCCGCAGAGACAGTAAAGCAAAAGAGGGCTCCGAGAACCCACAGTCT	1685
Qy	1621	tccgagtgccagcccgccgacgagatgaagaaagaaagaggtctccgaaccacagctct	1680
Db	1686	CAGTGA 1691	
Qy	1681	cagtga 1686	
RESULT	2		
LOCUS	AF051427	2011 bp	mRNA PRI 19-MAR-1998
DEFINITION	Homo sapiens	estrogen receptor beta mRNA, complete cds.	
ACCESSION	AF051427		
NID	92970563		
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;		
	Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE	1 (bases 1 to 2011)		
AUTHORS	Moore,J.T., McKee,D.D., Moore,L.B., Jones,S.A., Su,J.-L.,		
	Slentz-Kesler,K., Horne,E.L., Kilweier,S.A., Lehmann,J.M. and		
	Willson,T.M.		
TITLE	Cloning, ligand Binding and Functional Activity of Human Estrogen		
	Receptor beta isoforms		
JOURNAL	Unpublished		

RESULT	2			
LOCUS	AF051427	2011 bp	mRNA	PRI 19-MAR-1998
DEFINITION	Homo sapiens estrogen receptor beta mRNA, complete cds.			
ACCESSION	AF051427			
NUM	92970563			
KEYWORDS	human.			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 2011)			
AUTHORS	Moore,J.T., McKee,D.D., Moore,L.B., Jones,S.A., Su,J.-L., Slettz-Kesler,K., Horne,E.L., Kilewer,S.A., Lehmann,J.M. and Willson,T.M.			
TITLE	Cloning, Ligand Binding and Functional Activity of Human Estrogen Receptor beta isoforms			
JOURNAL	Unpublished			

Db 2007 AGTGA 2011
|||||
QY 1682 agtga 1686

RESULT 3
LOCUS AF051428 2041 bp mRNA PRI 17-MAR-1998
DEFINITION Homo sapiens estrogen receptor beta2 splice variant mRNA, complete cds.
ACCESSION AF051428
NID 92961558
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 2041)
AUTHORS Moore, J.T., McKee, D.D., Moore, L.B., Jones, S.A., Su, J.-L.,
Horne, E.L., Klierer, S.A., Lehmann, J.M. and Willson, T.M.
TITLE Cloning, Ligand Binding and Functional Activity of Human Estrogen
Receptor beta Isoforms
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2041)
AUTHORS Moore, J.T., McKee, D.D., Moore, L.B., Jones, S.A., Su, J.-L.,
Horne, E.L., Klierer, S.A., Lehmann, J.M. and Willson, T.M.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-1998) Molecular Sciences, Glaxo Wellcome, 5 Moore
Drive, RTP, NC 27709, USA
FEATURES
source Location/Qualifiers
1. 2041
/organism="Homo sapiens"
/db_xref="taxon:9606"
/rname="beta2"
/rnote="beta2 splice variant"
/codon_start=1
/product="estrogen receptor beta2 splice variant"
/db_xref="PID:g2961559"
/translation="MDIKNSPSSLSNPSSTVCSQSIIPLEHGSITIPSSIVDSHEHP
AMTFSPAVMYNSIPSNVTNLEGGPGRTSPNVLMPGLSLVHRLSHLYAP
QKSPWCBSARSLFHTLPVNRFTLKRKVSNGNCAVTPGSLPGLVAVHRLSHLYAP
YGVSCGCKAFKFRSIOGHNDYICPATNOCTIKNRKSCQALRKCYGVAVKGG
SRRECGYRLVRRRSADDEOLHCAGKAKRSGHAPRYRELLIDSLPOLYTLLEAE
PPHYLISRPSPFTEBAMMSLTILADKELVHMTSMKKTIGFELSLFDVRLLESC
WMEVLMMGLMWRSIDHPGKLIAPDLVLDROGKCVSGILEIFMLLATISREFELU
QKHEYLCKAMILNLSNMYPLVATODADSRRKLAHLINATDILVWYVIAKSGIS500
OSMRILANLMLSHVRAAPRAKRAKASQTLTSFGMKMETLLPEATMEO"

BASE COUNT 505 a 532 c 544 g 460 t

Query Match 88.7%; Score 1496; DB 21; Length 2041;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 327 AGCCATTACTGCGCCAGCAATCTTTGAGAACATTAATATACCTTTGGCTTTCTT 386
|||||
QY 2 agccattactgcccagcaatcttggagaacattatgaaccttggctcttctt 61

Db 387 GCAAGGTGTTTCTCAGCTGTATCTCAAGCATGATATAAAACTCACCATCTAGCC 446
|||||
QY 62 gcaaggtgttttctcagctgtatctcaagacatgatatataaaaccaccatctagcc 121

Db 447 TTAATTCCTTCCTCCTACCACTGCAATCATCTTACCCCTGGAGCAGGCTCCA 506
|||||
QY 122 ttaatctctctctcctcaacacgagtcacatcttaccctctggagcagcgctcca 181

Db 507 TATAATACCTTCCTCCTATGTAGACAGCAGCATGATATCCACCATGACATCTATA 566
|||||
QY 182 tataataacctctccctcatgtagacaagccacatgaatcaccatctgaattctata 241

Db 567 GCCCTGCTGTGATGATATACAGCATTCAGCAATGTCTACTACTTGAAGGTGGCCGG 626
|||||
QY 242 gccctgctgtgataatatacagcatctccagcaatgctactaacttggaaggtggccgg 301

Db 627 GTCCGACAGACCAACAAACCCAAATGTGTTGGCCCAACACTGGCCACCTTTCCTTTAG 686
|||||
QY 302 gtccgacagaccacaagcccaaatgtgtgtgccaacaccctggcaccttctccttag 361

Db 687 TGGTCATCGCCAGTGTATCATCTGTATGGCAACCTCAAAAGATCCCTGGGTGTAG 746
|||||
QY 362 tggctcatcgccagttatcatctgtatggcaacctcaaaagatccctgggtgtag 421

Db 747 CAAGATCGCTAGAACACACCTTACCTGTAAACAGAGACACTGTAAAGAGAGTTAGTG 806
|||||
QY 422 caagatcgctagaaacaaccttacctgttaacagagagacactgaaagaggtgtgtg 481

Db 807 GGAACCGTTGGCCAGCCCTGTTACTGGTCCAGGTTCACAGAGAGATCTTCATCTGCG 866
|||||
QY 482 ggaacctgtggccaagccctgttactgtgtccaaggttcaaaagagatgctcaacttcg 541

Db 867 CTGTCTGACGAGATTAGCATCGGAGATATCACTATGAGTGGTGGTGGTGAAGATGTA 926
|||||
QY 542 ctgtctgacgagattagcatcgagatatacactatgaggtctgtgtgtaagagatgta 601

Db 927 AGGCTTTTAAAGAGACATTCAGAGACATTAATGATTAATTTGTCACACTCAAAATC 986
|||||
QY 602 aggcctttttaaagaagacatctcaagacatgatatgattatgttccagctataac 661

Db 987 AGTGTACATGATTAATAACCGCGCAAGAGCTGCCAGGCTGCCAGCTTGGAGTGT 1046
|||||
QY 662 agtgtacatgataataaaacccgagcaagatctcagagctcgcagacttcgagatgt 721

Db 1047 ACGAAGTGGAGATGTGAGTGTGGCTCCCGGAGAGAGATGTGGTACCGCTTGTGC 1106
|||||
QY 722 acgaagtggagatgtgtgaggtgtgtctccgagagagagatgtgtgtgtaacgcttgc 781

Db 1107 GGAGACAGAGAGAGTCCGACGACGAGCTGACACTGTGCGCGCAAGAGAGAGTGGCG 1166
|||||
QY 782 ggagacagagagatgtgtgagagagagagagagagagagagagagagagagagag 841

Db 1167 GCCACGCGCCCGGAGTGGGAGCTGCTGTGACAGCCCTGATGCCCGAGCAGCTAGTGC 1226
|||||
QY 842 gccacgccccgaggtgcgagagctgtgtgagagagagagagagagagagagagagag 901

Db 1227 TCACCTCTCTGAGAGCGAGCGGCCCATGTCGATACGCGGCCGAGTGGCGCTTCA 1286
|||||
QY 902 tcacctctctgagagctgtgagagagagagagagagagagagagagagagagagag 961

Db 1287 CCGAGGCTCCATATGATGTGCTCCAGCAAGTGTGGCGCAAGAGAGTGTATACATGA 1346
|||||
QY 962 ccgaggtcccatatgatgtgtccctgacaaaggtgtgagagagagagagagagagag 1021

Db 1347 TCAGCTGGCCCAAGAGATTCGCGCTTGTGAGACTCAGCTGTTCGACCAAGTGGCGC 1406
|||||
QY 1022 tcagctggcccaagaagagatcccgcttgtgtgagagagagagagagagagagagagag 1081

Db 1407 TCTTGGAGAGCTGTGAGTGGAGAGTGAATGAGAGGCGTATGAGGCGTCATTTGACC 1466
|||||
QY 1082 tcttggagagctgtgtgag 1141

Db 1467 ACCCGGCAAGCTATCTTCTCCAGATCTTGTGGAAGAGAGAGAGAGAGAGAGAGAG 1526
|||||
QY 1142 acccggcaagctatctcttctccagatctgttctggaagagagagagagagagagag 1201

Db 1527 TAGAAGGATTCGTGAAATCTTTGACATGCTCTGGCAACTACTTCAAGGTTTGAGAGT 1586
|||||
QY 1202 tagaaggaattctgaaatcttgaacatgctctggcaactacttcaaggttctgagagt 1261

Db 1587 TAAAACTCAACACAAGAATATCTCTGTGTCAAGGCGCATATCTGTCAATTCCAGTA 1646
|||||
QY 1262 taaactcacaacaagaatatactctgtgtcaagagccatgatactgtcattccagta 1321

Db 1647 TGTACCTCTGTGTACAGAGCAGCCAGAGATGCTGACAGAGCGGGAAGGTGACTACTTGC 1706
|||||
QY 1322 tgtacctctgtgtacag 1381

Db 1707 TGAACGCCGTGACCGAGTGTGTTGGTGGATGGCCAGAGGCGCATCTCTCCAGC 1766

Db 1819 AGCATTCATGCCCTGGCTACCTCTGTCCTGTCCTCCAGTCAGCATGCGAG 1876
|||||
Qy 1442 agcaatcatgcgcctgctgaacctctgatgctctgtcccaagcagcagcag 1499
|||||

RESULT 5 AB006589 3593 bp mRNA PRI 28-JUL-1998
LOCUS Homo sapiens mRNA for estrogen receptor beta cx, complete cds.
ACCESSION AB006589
NID 93345671
KEYWORDS estrogen receptor beta cx.
SOURCE Homo sapiens CDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (sites)
Ogawa,S., Inoue,S., Watanabe,T., Orimo,A., Hosoi,T., Ouchi,Y. and
Muraiatsu,M. (1997) to the DBJ/EMBL/GenBank databases. Sumito
Submitted (13-AUG-1997) to the DBJ/EMBL/GenBank databases. Sumito
Ogawa, Saitama Medical School, Department of 2nd Biochemistry; 38
Morohongo, Moroyama, Iruma-gun, Saitama 350-0495, Japan
(E-mail:suogawa@saitama-med.ac.jp, Tel:81-492-76-1490,
Fax:81-492-94-9751)

FEATURES
source location/Qualifiers
1..3593
/organism="Homo sapiens"
/db_xref="taxon:9606"
1276..2763
/codon_start=1
/product="estrogen receptor beta cx"
/db_xref="PID:d1032928"
/db_xref="g3345672"
/translation="MDIKNSPSSLNSPSSYNSQSILPHEGSIYIPSSYVDSHHEYP
AMTFYSPAVNATSIPIGNVNIIEGPGRTSPVMTPTGHSPLVHNSLHLYAE
OQSPCEARSLEHTLIPVNETLTKRKYSGRKASPVLPGRKRAHCAQSPYASGYH
YGVMSCEGKAPKRSIOGHNDYICPATNOCITDKRRKSCAPRKRYEYGMVGC
SRREKGYRUVRRORSADPOLHRCARAKRSVGHAPVRELLDALSPEQVLTLEAE
PHEVLTMRSDPTEASMMSLTKLKDKEVHTSAKPIGVEVLSLTDQVRLBEC
MMEVLMGMWRSIDHPGLIFAPDLVLDRECKVEGLEIFDMLAATSRRELKL
QKREYLCVAMAILNLSMYPVLTATQDADSSRKLALHVAVDLWVATLAKSGISSQ
QSMRLANLMLLSHVHRAERAKESQRTLTFSGMKMETLPEATMEQ"

BASE COUNT 971 a 811 c 957 g 854 t
ORIGIN

Query Match 88.7%; Score 1495; DB 21; Length 3593;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1497; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1183 CAGCATTTACTTGGCCAGAACTTTGAGACATTAATAGCTTTGCTCTTCT 1242
|||||
Qy 1 caagccataactctgcccacgaatcttggaaattataatgaccttctgctctct 60
|||||

Db 1243 TGAAGGTTGTTCTAGCTGTATCTCAAGACATGATATAAAATCAGCATTCAG 1302
|||||
Qy 61 tgcgaagtgcttctccagcgcgtatcccaagacatgatatataaaactcaactcagc 120
|||||

Db 1303 CTAAATCTCTCTCTCTCAACATGCAATCAATCAATCTTACCCCTGGAGCAGGCTCC 1362
|||||
Qy 121 ctaaatctctctctctctcaactgagtaactcaatcttaaccctggagcaagctcc 180
|||||

Db 1363 ATATACCTACTTCTCTCTATGATAGACAGCAGCATTAATCAAGCATGACTTAT 1422
|||||
Qy 181 atataactactctctctctctatgttagacagcacaatgaataatcaagcagacttctat 240
|||||

Db 1423 AGCCCTGCTGATGATTAATCAGATCCAGCAATGATCACTAATGGAAGTGGGCT 1482
|||||
Qy 241 agccctgctgtgtagaattacaagatcccaagcaatgtcactaacttggaagtgaggct 300
|||||

Db 1483 GGTGGGACAGCAAGCCCAAAATGTTGTGGCCAAACCTTGGGACCTTTCTCTTTA 1542
|||||
Qy 301 ggtcgcagacacacaagcccaaatgtgtgtgtccaaacacactggaacctctctctta 360
|||||

Db 1543 GTGCTGCATCGGCAGTTATACATCTGTTGGGGAACCTCAAAAGTCCCTGGGTGTA 1602
|||||
Qy 361 gtgtccatcgcagagttatcaatcattgtagcggaaacctcaaaagagctccgtgtgaa 420
|||||

Db 1603 GCAGATCGCTAGAACACACACTTACCTGTAACAGAGACACTGTAAGAGAGTTACT 1662
|||||
Qy 421 gcaagtcgctatgaatacaacaccttactctgtaaacagagaaacatgaaagaggttagt 480
|||||

Db 1663 GGGAACCTGTGGCCAGCCCTGTACTGTCCAGTTCAAAGTTCAAAGGATCTCACTTTC 1722
|||||
Qy 481 gggaaacctgtgcagccctgttactgttcaagttcaagagagatgtcaactctgc 540
|||||

Db 1723 GCTGCTGAGGATTAACCATCGGATTCCTACTATGAGTGTGTCGTGTAAGGATGT 1782
|||||
Qy 541 gctgtctgagagatcagcatcggatatacactatgagctgtgctgtgtgaagatgt 600
|||||

Db 1783 AAGGCTTTTAAAGAGCATTCAGGACATTAATGATTAATTTGTCACGTACAAAT 1842
|||||
Qy 601 aaggtctttttaaagaagcatcaagacataatgatatattgtccagcttaacaat 660
|||||

Db 1843 CAGTTACATCGATTAATAAACGGCGCAAGAGCTGCCAGGCTCCGACTTGGAAAGT 1902
|||||
Qy 661 caggtacaaatcgatataaaacgycgcaagagctgcagagcctgcagacttcggaagtgt 720
|||||

Db 1903 TACGAGTGGGAATGGTGAAGTGTGCTCCCGAGAGAGAGATGTTGGGATACGGCTGTG 1962
|||||
Qy 721 taagaagtggaatggtgtgaagtggtgtcccggaagagagatggtgtaccgctctgtg 780
|||||

Db 1963 CGGAGACAGAGAAGTCCCGACGACGAGCTGACTGTGCCGGAAGGCCAAGAGAAGTGGC 2022
|||||
Qy 781 cggagcagaagaagtgtcgcagagcagctgcactgtgcccgaagccaagagaagtgtgc 840
|||||

Db 2023 GGGCAGCGCGCCGAGTGGGGAGCTGCTGTGTGAGCGCCCTGAGCCCGACAGACTAGTG 2082
|||||
Qy 841 ggcacagcgcccgagtgctgagagctgtgagcctctgagcccgagcagctagtg 900
|||||

Db 2083 CTACCTCTCTGAGGCTGAGCGCGCCCAATGCTGATACGCGCCCGAGTGGCGCTTC 2142
|||||
Qy 901 ctacacctctctgagagctgtgagccgacatgtgtatcaagccgccccagctgtgccccctc 960
|||||

Db 2143 ACCGAGGCTTCATGATGATGTCCTGACCAAGTTGGCCAGAGAGTGTGATACATG 2202
|||||
Qy 961 accgaggtccctcatgtatgtgtccctcgacaagtgtgcccgaagagatgtgtacacatg 1020
|||||

Db 2203 ATCAGCTGGGCAAGAAAGATTCGGGCTTTGTGAGCTACGCTGTTCAACAAAGTACGG 2262
|||||
Qy 1021 atcagctggcccaagaagatctccgctgtgtgagctcagcctgttcgacccaagtgcg 1080
|||||

Db 2263 CTCTGGAGAGCGTGGATGAGAGGTAAATGATGGGGGTGATGGGGCTCAATTGAC 2322
|||||
Qy 1081 ctcttggaagagctgtggaagagtgtaatgaaagagagagagagagagagagagagagag 1140
|||||

Db 2323 CACCCCGCAAGCTCATCTTTGCTCCAGATCTTGTCTTGAGCAGGATGAGGGAAATGC 2382
|||||
Qy 1141 caccocggcaagctcatcttctgtccagatctgtctcgaagagagatgagggaaatgc 1200
|||||

Db 2383 GTAGAAAGAAATTTGGAATCTTTGACATGCTCTCTGGCAACTTCAAGTTTCAGAG 2442
|||||
Qy 1201 gtagaaggaattctggaatcttgcacatgtctctcgtgcaactactcaaggttctcgaag 1260
|||||

Db 2443 TTAAGCTCCAAACAAGAAATATCTGTGTGAAGGCAATGATCTCGCATTAATCCAGT 2502
|||||
Qy 1261 cttaaactccaacaagaataatctctgtgtcaagagcagtgatccctgcaatctcagct 1320
|||||

Db 2503 ATGTACCTCTGTGTACAGGACCCAGAGATGTGAGACAGACGCCGAGAGTGGCTCACTTG 2562
|||||
Qy 1321 atgtacctctgtgtcacaagagaccagagatgtctgacaagcagcagagctgtgctcactgt 1380
|||||

Db	2563	CTGAACGCCGAGCGATGCTTTGGTTGGGTGATGTCGAAGACGGGATATCTCCAG	2622
Oy	1381	ctgaacgcgcggaacgaagcttggttgggtgatgtccaagatccacag	1440
Db	2623	CAGCAATCCATGCGGCTGGCTTAACCTCTGATGCTCTGTCCACGTCAGGATCGCAG	2681
Oy	1441	cagcaatccatgcgccgtgctaactctgtgtctctgcgtccacgctcagcagcag	1499
RESULT	6	standard; DNA; HDN; 1460 BP.	
ID	A61586		
AC	A61586;		
NI	e1262494		
DT	09-MAR-1998 (Rel. 54, Created)		
DT	09-MAR-1998 (Rel. 54, Last updated, Version 0)		
DE	Sequence 4 from Patent WO9709348.		
OS	Homo sapiens (human)		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates;		
OC	Cararrhini; Homnidae; Homo.		
RA	[1]		
RA	Kulper G.G., Enmark E., Gustafsson J.;		
RT	"ORPHAN RECEPTOR";		
RL	Patent number WO9709348-A/4, 13-MAR-1997.		
RL	KAROBIO AB (SE)		
FT	key	Location/Qualifiers	
FT	source	1. 1460	
FT		/organism="Homo sapiens"	
SO	Sequence 1460 BP; 352 A; 379 C; 411 G; 318 T; 0 other;		
Query Match		86.1%; Score 1452; DB 9; Length 1460;	
Best Local Similarity		99.8%; Pred. No. 0.00e+00;	
Matches 1455; Conservative		0; Mismatches 3; Indels 0; Gaps 0;	
Db	3	ATGACATCTACAGTCCGTGCTGTGATGAATACAGCAATCCAGCAATGTCATACTTG	62
Oy	229	atgacatctctagctgacctgctggtatgatattacagcatctccagcaatgltacacttg	288
Db	63	GAAGGTGGCCCTGTGTCGGCAGACACAAAGCCCAATGTGTGTGGCCACACCTGGGCAC	122
Oy	289	gaagtggtgctgtgtgcgcagaccacaagcccaatgtgtgtgtgccaaacctggcac	348
Db	123	CTTTCCTCTTAATGATGTCATCCGCAATGATTCACATCTGATGCGGAACCTCAAGAAGT	182
Oy	349	ctttccctcttaagtggtccatcgcagattacacatctgtatgcygaacctcaaaagt	408
Db	183	CCCGGTGTGAAGCAAGATGCTTGAACACACCTTACTCTTAACAGAGACACTGA	242
Oy	409	cccggtgtggaagcaagatcgctgaagaacaaccttacctgaacagagagacactgaaa	468
Db	243	AGCAAGGTACTGGAACCCGTGGCCAGCCCTGTACTGCTCCAGGTTCAAAAGAGAT	302
Oy	469	agcaaggtatgtaggaacacgtgtgcgcagccctgttactgtgtccaggttcaagaaggtat	528
Db	303	GCTCATCTTCGCGCTGTCTGCAGCGATTACGATCGGATATCACTATGAGTGTGTCG	362
Oy	529	gctcactcttcgctgtctgcagcgattacgcacatcgcgatatcactatggtctgtcg	588
Db	363	TGTGAAGATTAAGGCTTTTAAAAAGAGCATTCAGAGACTAATGATTATTTGT	422
Oy	589	tgtgaagatgtaagcctctttttaaagaagcatccaagagacataatgattatttgt	648
Db	423	CCAGCTACAATTCAGTGTACATGATGATAAAAACGGCGCAGACACTGCGAGCCTGGCGA	482
Oy	649	ccagctacaatctcgtgtacacatcgataaaaaacggcgcaagagctgccaggtcctgcga	708
Db	483	CTTCGGAAGTGTACGAAGTGGAGTAAGTGTGATGTTGGCTCCCGAAGAGACAGATGTGG	542
Oy	709	cttcggaagtgttacgaagtgtggaatggttgaagtgtgtctccggagagagagtgtggg	768
Db	543	TACCGCTTTGTGGCGAGACAGAAAGTCCGACAGACAGCTGTGCTGCCGGAAGGCC	602

Oy	769	tacggccttgygcgagacagagaagtgcgcgaagcagactgtcacctgtccgcgaagcc	82
Db	603	AAGAGAATGGCGGACCAACGCCGCCGAGTGCGGAGCTGTCTGACAGCCCTTGAGCCC	662
Oy	829	aagaagaatgygcgcacacgcgcccgagtgcyggaagctgtctgtgcagcccttgagccc	888
Db	663	GAGCAGCTAAGTCCTACCCTTCTTGAGAGCTTAGAGCCGCCCATGTGCTGATCAGCCGCC	722
Oy	889	gagcagctagtgctcaacctctctgtgaagcttagagccgcgcccatgtgtctgatcaaccgc	948
Db	723	AGTGGCCCTTCACCGAGAGGCTCCATCATGATGTCCTGTGACAAAGTTGGCGCACAAAG	782
Oy	949	agtgcgcccccttcacccagagctccatgatgatgtcccttgaccaagttgcgcgaagaag	1000
Db	783	TTTGTCACATGATCAGCTGGGCCCAAGATTCCCGCTTGTGTGAAGCTCAGACTGTTC	842
Oy	1009	ttgtgacaatgatcagctggccccaagaagattcccgcttftgysagctcagcctgttc	1060
Db	843	GACCAAGTGC GGCTTGTGAGAGCTGTTGGATGGAGGTGTTAATGATGGGGCTGATGTGG	902
Oy	1069	gaccaaaagtcggtcttgyaagactgttgtatgtaagggtgtttaaatgataaggggctgagt	1122
Db	903	CGCTCAATTGACCAACCCCGGAGACTCATCTTTCCTCCAGATCTGTGTTCTGGACAGGAT	966
Oy	1129	cgtccaattgtcacaccccggaagatcalctctgtgtccaaatctgttccggaagggat	1188
Db	963	GAGGGAAATCTGTAGAGAAGTAATCTGGAATCTTTGACATGCTCCGCGCACTACTTCA	1020
Oy	1189	gaggggaaatgtcgtagaagaatctcttgyaaaactcttgaatgtccctgcgaactactca	1240
Db	1023	AGGTTTGAGAGGTTAAAACCTCCAACACAAGAAATATCTGTGTGTCAAGGCCATGATCTG	1080
Oy	1249	aggttcgaggtgttaaactccaacacaagaatatctctgtgtcaagcgaatgtatctg	1300
Db	1083	CTCAATTCOAATATGTCOCCTCTGTGTGACAGCGAACCCAGAGATGCTGACACGCGGAAG	1140
Oy	1309	ctcaatctccagatgatccctctgtgtcacagcgaccacaagatgtgtgacagcagccgaag	1360
Db	1143	CTGGCTACTTGTGCGAACGCCGCTGACCGATGCTTTGGTTTGGGTGATGGCAAGCGGC	1200
Oy	1369	ctgtgtcaactgtcgcgaagccggtgtaccgaatgtgttgggtgtgtgttgcgaagagcgc	1420
Db	1203	ATCTCTCCACGACAGCAATCCATGCGCTGTGCTAACCTCTGTATGCTGTCTGCCAGCTC	1260
Oy	1429	atctctccacagcagcaatccatgtgcgtgtctaactctgtatgtctctgtccacgctc	1480
Db	1263	AGGCAATGGAGTAACAGGGGATGGGAACATCTGCTCAACATGAAAGTGC AAAATGTGTC	1320
Oy	1489	aggatctgagatlaacaagggatgaaacatctgtctaacatgaagtgcataaaatgtgttc	1540
Db	1333	CCAGTGTATGACCTGCTGCTGGAATCTGTAATCCCACTGCTCCCGGCTGCAAGTCC	1380
Oy	1549	ccagtgtatgtacctgtcgtcgtggaatgtctgaatgtccaaatgtctgtgcggtgtcaagtc	1600
Db	1383	TTCATCAGCGGGTCCGAGTGCAGCCCGGACAGAGGACAGTAAAAAGCAAAGAGGCTCCAG	1440
Oy	1609	ttcatcaaggggtctcgtggtgagccgcgcagagtaagcaaggaaggggtctccag	1660
Db	1443	AACTCAAGTCTCAGTGA 1460	
Oy	1669	aaccacagttctcagtga 1686	
RESULT LOCUS	7 HSRNAERB 1560 bp RNA PRI	05-SEP-1996	
DEFINITION	H.sapiens mRNA for estrogen receptor.		
ACCESSION	X99101		
NID	g1518262		
KEYWORDS	estrogen receptor.		
SOURCE	human.		
ORGANISM	Homo sapiens Eukaryota; Eutheria; Primates; Catarrhini; Homnidae; Homo.		

Accession	NID	Keywords	Source	Organism
AJ002602	g3077648	estrogen receptor beta.	Norway rat.	Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrate; Mammalia; Eutheria;				
Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
1 (bases 1 to 1650)				
Aldridge, T.C.				
Tissue specific responses to estrogen: an explanation based on differential activation of multiple estrogen receptors with different estrogen response elements				
Unpublished				
JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL
Submitted (17-APR-1997) Aldridge T.C., Molecular Endocrinology Group, ZENECA Central Toxicology Laboratory, Alderley Park, Macclesfield, Cheshire SK10 4TJ, ENGLAND				
(bases 1 to 1650)				
Aldridge, T.C.				
Direct Submission				
Submitted (18-NOV-1997) Aldridge T.C., Molecular Endocrinology Group, ZENECA Central Toxicology Laboratory, Alderley Park, Macclesfield, Cheshire SK10 4TJ, ENGLAND				
revised by [3]				
location/Qualifiers				
1..1650				
/organism="Rattus norvegicus"				
/strain="Wistar"				
/db_xref="taxon:10116"				
/clone="pcDNA_Erbeta"				
/tissue_type="Prostate"				
1..1650				
/gene="Erbeta"				
1..1650				
/gene="Erbeta"				
/note="The clone was isolated by RT-PCR from total RNA"				
/codon_start=1				
/evidence="experimental"				
/product="estrogen receptor beta"				
/db_xref="PDB:el287358"				
/db_xref="PID:g3077649"				
/translation="MSICTSHKEFSOLFSEDEIKNPSSLSPPASYNCOSILPLGHSYIPSSVYDNRHEYSAMTFSPAVMNVSVPGTSNLDGPVRQSTSPNYMPISGHGIPIATHOSSLILYAEPOKSPMCACRALEPHTPVARETLKRSLSSCASPTSPN AKRAHFCAYCSDIASGYHYHWPCBEGKAFFKRSIGCHNDYTDPATNOCTIDNRKK SCQCLRLKCTCEVGWVKSGSRERCGRYIVARRQSSBOVCLSKARNGHARVEE LLSLTPEDOLVLLLEAEPENPVLRSMSPFTASMMSLTKLADKLVIIMGMARK IPGEVELSILDQVRLIESCMWEVLWGIMSGSIDHPGLIAPDLVIDRDGCKVEGII LEIDMLLATSRRELKHOKKEYLKAYMLLNSSMYPLASANOEAESSKLTHLNL AVPMALVMVIAKSISISSOQDVRLANLMLLSHRVSHSNKMHELLMKCNVPYTD LLEMLNHNTLRGPKPSISGSSESTDSKNKSSONLOSO"				
BASE COUNT	415 a	443 c	432 g	360 t
ORIGIN				
Query Match	67.6%	Score 1140:	DB 22:	Length 1650:
Best Local Similarity	85.2%:	Pred. No. 0.00e+00:		
Matches 1399:	Conservative 0:	Mismatches 240:	Indels 3:	Gaps 2:
Dd	10	TGTACTCTTCACAGAGATTTCAGCTGCAGCCCTCGAAGACATGGAGATCAA	69	
Oy	49	tgtgcctcttccttggaagtggtttcctcagctgtatc-tc--aagacatgatataaa	105	
Dd	70	AACCTACCGTGAGCGCTAGTTCCCTCTGCTTCTCTATACTGTAGCCAGTCACTACC	129	
Oy	106	aactacaacctcagcttaattctcttccctcaactcagcatcactccttaccc	165	
Dd	130	CTGGAGCACGGGCCCATCTACATCCCTTCCTCTAGTGAACAACCGCATGATATCA	189	
Oy	166	ctggagcaagcgctcataatacactcctccctctgttagaacaccacccaagaatatcca	225	
Dd	190	GCTATGACATTCCTACAGTCTCTGATGAATCACTACAGTGTCCGGGAGCACCACTAAC	249	

QY	226	gcatgacattctatagccctctgtatgaattacagcatctccagcaatgtaactaac	285
Db	250	ctggacagctgggacctgtgccacagacacaaagcccaattgcttatggccaaactttctggg	309
QY	286	ttggaagttggtgacctgtctcgtgcagaccacaaagcccaattgtctgtgccaacaccttgg	345
Db	310	caactgttcccttttagagaccattggccatgcatnccgtctcttatgcacaaactctaaag	369
QY	346	caacttctcccttaagtggtccacacgcgaattcacatctgtatctgtacggaacctcaag	405
Db	370	acttccttggcttgaaagcaagatcactagacacacacttaactcttaataacagagacactg	429
QY	406	agtcctctgttgaagcaagatcgctgaacacacacttaactctgtaaacgaagagacactg	465
Db	430	aagaggaactttagtggagcagttgtgcacgacctgttatagtccaagccaaagag	489
QY	466	aaaagaaggttaagtbgnaaccgttgcgcagccctgtactgtccagttccaagaag	525
Db	490	gattgtcacttctggccgctgtcagaggaattatgcactnccggatnataccgaccttgg	549
QY	526	gctgtccactctctgcctgcgtctcgaagcattacgcacatcggatatacctaagtacatcgg	585
Db	550	ccatttgaagatgttaagacccttttttaaaagaagattcaagacatcatattatattc	609
QY	586	tcgtgtgaagatgttaagaccttltttaaagaagcatlccaaggaatataatattatatt	645
Db	610	ttgtccagccaggaattcagatgtaccatagacaaacccggcttaaaagctgccagacctgc	669
QY	646	tgtccaagctataaactagtgtaacatcgataaacaacggcgcaagctgtccagacctgc	705
Db	670	gcacttgcacagtgttttgaagtggaatgctcaatgtngganccagagagacaaagctgt	729
QY	706	cgactccggaagtgtctcgaagtgtggaatgtgaatctgtgctccggagagaagatgt	765
Db	730	gggtaccttatatagtcgagagcagagaaagtcttcagagagcagagcttacactgaccaa	789
QY	766	gggtacctgcctgtgtgcgagagcagaagaatgtccagagagatgtacactgtgcgcgaag	825
Db	790	gccaaagaaaacgctggcgacacacccgggcttaagaagctactgctgacacattgact	849
QY	826	gccaagaagaatgtgc	885
Db	850	ccagagcaactcggggcaccctctcttggaagctgaacacccaaatgctgtgctgagccgt	909
QY	886	cccgagcagctcagtgctcaaccctctcgtgaagctcgaagccgcacatgtctgtacagcgc	945
Db	910	cccaacaaagcccttcaccgagaccctccatattatgttccctcactaaagctgcgacaaag	969
QY	946	cccagtggtcccttcacacggagcctccaatgtatgttccctgcacagatgtgcgacaaag	1005
Db	970	gaactgtgtcactatgattggctggggccagaaaaatccctggcttcttgtagagctcagcctg	1025
QY	1006	gaagtgtacacatgatacagctgtggtccaaagaatctccggcttctgtgaagctcagcctg	1065
Db	1030	tttgacacaaagtcgggctctttagaaagctgctggaatggagagtgcttatgtgggactgatg	1085
QY	1066	ttcgaccaagtgccgacctcgtgaagctgtcttgaaatgtgaatgtgaatgtggtgcgtgag	1125
Db	1090	ttggacctctatcgcacacaccccggaacactattttgtgctcccaacctcttcttgacagag	1145
QY	1126	ttggcgtcaattgtgcaaccccggaagctcatcttctgtccacagatcttgtcttgacagag	1185
Db	1150	gattgagggaaatgcgttaagaagattcttgaatcttgaatctgtccctgtgcaactact	1245
QY	1186	ttcaagttccctgtgattaaaactccacacaaagaaatatactctgttccaagccatgac	1265
Db	1210	ttcaagttccctgtgattaaaactccacacaaagaaatatactctgttccaagccatgac	1265
QY	1246	ttcaagttctcgaagttcaaaactccacacaaagaaatatactctgttccaagccatgac	1305
Db	1270	ttcttccaactccagattatgaccttggcttcttgcaaacacagagacacaaagtgacggg	1325

QY	1306	ctgtcctaattccagatgatgtacccctctgtgtcaacagcgaccacgaagatgctgtcaagaagccgg	1365
Db	1330	AAGCTGACACACCTACTGAACGCGGTGTGACAGATGCGCTGTGTGGTGATGTGCAAGAGT	1389
QY	1366	aagctgtgcacttgctgtgaacgcggtgacgcgattgtgttggtgtatgttgcagaagac	1425
Db	1390	GGTATCTCTCTCCCGACGACGACTGATCCGACTGGCGCAACCTCCGATGCTCTTTCAC	1449
QY	1426	ggcactctctcccaagaacatcaatgactgactgtgaacctccgtatgctctcctccac	1485
Db	1450	GTCACGACATCATGATGAACGAGGCGATGGAATGCTGCTACACATGAAGTCAAAATGTG	1509
QY	1486	gtcaggaatcgaggtaaacaaaggctgtgaataatctgtcccaacatgaaagtgtcaaaaagt	1545
Db	1510	GTCCCGGTGTATGACCTCTGCTGTGAGATGCTGATGCTCAGACGCTTCGAGGTACAG	1569
QY	1546	gtccagagtgatgactgtgctgtggaagtgtcgtgaatgtccacgttcttcgcggtgtcag	1605
Db	1570	CCCTTAATCTGGGGGTCTGTGATGCTCAGGCTCAGAGGACAGTAAAGAAAGAGCTCC	1629
QY	1606	tctccatcaacagggtctcgaagtgcagcccgacagagacagtaaaagaagaaggtccc	1665
Db	1630	CAGAACCTACAGTCTCCAGTGA	1650
QY	1666	cagaaccacagctccagtgca	1686
RESULT	9		
LOCUS	RN057439	2555 bp	ROD 12-JUN-1996
DEFINITION	Rattus norvegicus estrogen receptor beta mRNA, complete cds.		
ACCESSION	U57439		
KEYWORDS	g1373280		
ORGANISM	Norway rat.		
SOURCE	Rattus norvegicus		
REFERENCE	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Rattus.		
AUTHORS	1 (bases 8 to 2555) Kulper, G.G., Enmark, E., Peltö-Huikko, M., Nilsson, S. and Gustafsson, J. A.		
TITLE	Cloning of a novel receptor expressed in rat prostate and ovary		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 93 (12), 5925-5930 (1996)		
REFERENCE	2 (bases 1 to 2555) Enmark, E.L.		
AUTHORS	Direct Submission		
JOURNAL	Submitted (03-MAR-1996) Eva L. Enmark, Center for Biotechnology, NOVUM, Karolinska, Institute, NOVUM, Huddinge, S-141 57, Sweden		
FEATURES	Location/Qualifiers		
SOURCE	1..2555		
	/organism="Rattus norvegicus"		
	/strain="Sprague Dawley"		
	/db_xref="taxon:10116"		
	/sex="male"		
	/tissue_type="prostate"		
	/note="nuclear receptor superfamily member"		
	/codon_start=1		
	/product="estrogen receptor beta"		
	/db_xref="pid:g1373281"		
	/translation="MTFTSPAVVMNYVPGSTSNLDGFPRLSTSPNVLWPTSGHLSPL		
	ATHGSSLLIYAEPOKSPWCEARSLSEHTLPVNERLTAKRLSGSSCASPTSPNAKRDH		
	PCPVCSDYASGHHGVMGCEGKAFKFKSIGDHNDICPANTOCTIDNRKRSQACOR		
	LKPCVGMVAKGSRREKRGYRIYVROSSSEOVHICSKARKNGGHAHPYDEILLSTLL		
	SPEDVLVTLLENPVNLVSRSPFTASMMSLITKLADELYHMGAKKIDPEVPE		
	LSLDQVRLLSCMEVLMVGLMWSIDHPKLIIFAPDLVLDREGKCEVSEITFDML		
	LIATSRRELKLOHKEVLMCYAMLLMSISMYPIASANOAESSRSKLTHLINAVYTDAL		
	VWVAKSGISSQOQSVRLANLIMLSHVRHLSNKGMEHLISMKCNVVPYDILLLEML		
	NAHTLRGYKSTISGSECSSTDSSKSKESQNLQSG"		
	726..923		
	misc.feature		
	/note="encodes DNA binding domain"		
BASE COUNT	630 a 652 c 660 g 613 t		

ORIGIN	Query Match	67.4%;	Score 1136;	DB 22;	Length 2555;
	Best Local Similarity	84.8%;	Pred. No. 0.00e+00;		
	Matches 1420;	Conservative	0;	Mismatches 248;	Indels 6; Gaps 5
Db	204	CCCATGAGCCTCTGGAACATATAG-TC-CANCTGTACTCTTCTCCAAAGAGTTTCT	261		
Qy	16	cccaagaaaccttgagaacataaagaaccttgcctctctcttcaaggttttct	75		
Db	262	CAGCTGCAGCCCTCTGGAAGACATGGAGATCAAAAACTCACGCTGAGGCTTAGTTC-CCT	320		
Qy	76	cagctgcgctatc-tc--aagaatgatatataaactcacatctagacttaattcct	132		
Db	321	GCTTCCTTAATACTGTAGCCATCTCATCTACCCCTGGAGCAGGCCCATATACATCCT	380		
Qy	133	tctctctcaactgagctcaactcctcttaccctctgcagcgtctcatataact	192		
Db	381	TCCTCTACGTATGAGAACCCGCATGAGTATTCAGTATGACATTTCTACAGTCCTGCTG	440		
Qy	193	tctctctcaatgataagacacataaatacagaacgaatctataagccctgcgtg	252		
Db	441	ATGACATCAGAGTGTCCCGGAGCAGCACAGTAACCTGGAGCGGTGGCTCGCATGAGC	500		
Qy	253	atgaaattacagatctcccaagcaatgctcaactacttggaagtgtgcctgcgcgaagacc	312		
Db	501	ACAAAGCCCAATGTGCTATGCGCAACTTCTGGGCACTGTCTCTTTAGCAACCATTTGC	560		
Qy	313	acaagcccaaatgtgtgtgtgcacaacactgtgcaccttctccttagtgcctatgcg	372		
Db	561	CATATCATGCTCTCTATGAGAACCTCAAAAAGATCTCTGTGGTGTGAGCAAGATCATTA	620		
Qy	373	cagttatcacatctgataagcgaaacctcaaaagagctccgtgtgtgaagcaagacgtcta	432		
Db	621	GAGCACACCTTACCTGTAACACAGAGACACACTGAGAGAGAAAGCTTAGTGGAGCATGTGT	680		
Qy	433	gaacacacacttaccctgtataacagagaagacacgtgaagaagatgtagtgggaaccgcttgc	492		
Db	681	GCCAGCCCTGTTACTAGTCCAAACGCAAGAGGAGTGTCACTTCTGCCCGCTGCAGC	740		
Qy	493	gcacagccctgttactgtgtcaggtgtccaagaagagatgtcactctgcgtctgcagc	552		
Db	741	GATTATGCAATCGGGATTCATTACGGCGCTTGGTGCATGGAAGATGTAAAGCCCTTTT	800		
Qy	553	gattacgcacatcgagatatcatcatgagctcgtgtcgtgtgaagagatgtgaagcccttctt	612		
Db	801	AAAAGAGATTTCAAGAGACATATATTATATCTTCAGCCACCAATACAGTACCTATA	860		
Qy	613	aaaagaagaattccaaggaataatataattatatttgcagctacaacaatcagtgatactc	672		
Db	861	GACAAAGAACCGGCGTAAAGCTGCCAGGCGCTGCCACTTCCGAAGTGTATGAAGTAGGA	920		
Qy	673	gataaaaacccgagcaagagctgtccagagcctgtcgactctgaaagtgttaagaatgtgga	732		
Db	921	ATGCTCAATGTGGATCCAGAGAGAACGGGTGTGGGTACCGTATAGTGGAGGAGAGAGA	980		
Qy	733	atgtataaagtgtggtcccccggagaagaatgtgtgtgtacgcgtctgtgcggagaagaaga	792		
Db	981	AGTTTACGAGAGAGATGACTGCTTGAGCAAGCAAGCAAGAAAGAGTGGAGCATCACCC	1040		
Qy	793	atgtgcgaagagagatctgacatctgtcgcgycgaaggtccaagaagaatgtgcgcagcgccc	852		
Db	1041	CGGGTAAAGAGAGTACTGCTGTGAGCACTTTGAGTTCAGAGCAACTGGTGTCTACCTCTGTG	1100		
Qy	853	cggagtgtcggagagctgtgtgtgtgaagccctgcgaagcccgagcagcttagtgttaacccctctg	912		
Db	1101	GAAGCTGAACCAACCAATGTGCTGGTGAAGCGCTCCAGCATCCTTACCGAGGCGCTTC	1160		
Qy	913	gaagctcgaagcgcgcccatgtgtcatcaacgcgcccaatgtgcgccttcaacgaagccttc	972		
Db	1161	ATGATATATTCCTCTCATAGCTGGGCGGACAAAGAAATGGTGCACATGATGTTGGCTGGGCG	1220		
Qy	973	atgatatgttccctgcgaaggtgtgcgcgaagagatgtgtatacaatgtctcagctcgggccc	1032		

Db	1221	AAGAAATCCCTGGCTTGTGTGAGCTACGCTGTGTGGACCAAGTCCGGCTTTAGAAAGC	1280
Qy	1033	aagaagatcccgccgttttgtagcctcagccctgttcgcaagfygcgctcttgagagc	1092
Db	1281	TGCGGATGGAGGTCCTAATGTTGGGACGTAGTGTGGGCTTCATCGACCAACCCGGGAG	1340
Qy	1093	tgtcgagaggggtagtaatgtagtgggctgtagtgcgtcctaattgacacccgggaag	1152
Db	1341	CTCATTTTCGCTCCCGACCTCGTTCTGTGACAGGAGATGAGGGGAGTGGTGAAGGATT	1400
Qy	1153	ctcaatcttgctccagatcttgctctgtagcaaggatgagggaaatgcytagaagagatc	1212
Db	1401	CTGGAATCTTTGACATCTCTCTGGCGACGCTGTAAGTTCCTGAGTTAAACTCCAG	1460
Qy	1213	ctggaatacttggacatctccttgccaactcttcaagagtttcgagagttlaaactc	1272
Db	1461	CACAAGAGTATCTGTGTGAGAGGCCATGTATCCCTCCATCTCCAGTATGATACCTCTTG	1520
Qy	1273	cacaagaataatctcttgtagcaaggccatgctccgcaatccagatgtagtacccttgc	1332
Db	1521	GCTTCTGCAAAACAGAGAGGAGAAAGTAGCCGGAAGCTGACACACACTGATGACGCGGTG	1580
Qy	1333	gtcacagagaccacgaatgctgtagcaagcagccggaagctgctcaactgtgtgaagcggtg	1392
Db	1581	ACAATAGCCCTGGTGTGGGTGATTTCCGAAGTGTATCTCTCTCCAGACGATCAGTC	1640
Qy	1393	accgaatgtttggtttggtttgtagttagcgaagcgggcatctctccacgagcaatcatg	1452
Db	1641	CGACGTGGCAACCTCTATGCTCTTCTCTACGTCAGCAGCATCAGTAACAGAGGATG	1700
Qy	1453	cgctctggttaactccctcgtatgctctgtcccaagtcagagcatgagataaaggagatg	1512
Db	1701	GAACATCTGCTCAGCATGAGTGCAGAAAATGTGGTCCGGTGTATGACCTGCTGGAG	1760
Qy	1513	gaacatctgccaacatgaatgcaaaaatgtgtgtccagtgtagtaacctgctcgag	1572
Db	1761	ATGCGATATGCTCAGCCTTCAGAGGCTACAAAGTCTCAATCTGGGGTCTGAGTCAGC	1820
Qy	1573	atcgctgaatggccaagtagctctcgagggggagcaagttctctcaaggggcttcgaagtcagc	1632
Db	1821	TCAACAGAGCAGTAAAGAACAAAGAGAGCTCCGAGAACTATACGTCACTGTA	1874
Qy	1633	ccggcagagggagcagtaaaagcaaaagagggctcccaagaaccacagctctcagtg	1686
RESULT 10			
ID	A61583	standard; DNA; ROD; 2568 BP.	
AC	A61583;		
NI	e1262491		
DT	09-MAR-1998 (Rel. 54, Created)		
DT	09-MAR-1998 (Rel. 54, Last updated, Version 0)		
DE	Sequence 1 from Patent WO9709348.		
FW			
CC	Rattus rattus (black rat)		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia;		
CC	Sciurognathi; Muridae; Murinae; Rattus.		
RN	[1]		
RA	Kulper G.G., Enmark E., Gustafsson J.;		
RT	"ORPHAN RECEPTOR";		
RL	Patent number WO9709348-A/1, 13-MAR-1997.		
RL	KAROBIO AB (SE).		
FH	Key	Location/Qualifiers	
FT	source	1..2568	
FT		/organism="Rattus rattus"	
FT	Sequence 2568 BP; 634 A; 654 C; 663 G; 617 T; 0 other;		
Query Match 67.4%; Score 1136; DB 9; Length 2568;			
Best Local Similarity 84.8%; Ped. No. 0.00e+00;			
Matches 1420; Conservative 0; Mismatches 248; Indels 6; Gaps 5			
Db	211	CCCATGAGCTCTGTGAGACATAATG-TC-CATCTGTACTCTTCTCACAGAGATTTC	268

QY	16	cccaagaa	tctttt	agaacatt	ataatga	tccttt	tgccct	ctcttc	gtcgaag	ttgtctc	75
Db	269	CAGCTGCAC	CCCTCTG	AAGACAT	GTGAGAT	CAAAAC	TCACCG	TGCAAGC	CTTATG	TTC-CCT	327
QY	76	cagctgc	atc-tc-	-aaga	catg	atataa	aaactca	acatct	agact	tattctct	132
Db	328	GCTTCCT	TAACT	TAGCCAG	TCCATCT	TACCCCT	TGAGAC	AGGCC	CCATCT	TAATCCT	387
QY	133	tctctct	caact	ctg	cag	tcaat	catctt	ccccct	tgag	ca	992
Db	388	TCCTCT	CTAG	AGAC	CCCGCAT	AGATG	ATATG	CACTAT	ATGAC	TGCTCT	447
QY	193	tctctct	at	gtaga	cacg	ccac	atga	atata	tcgag	caatct	252
Db	448	ATGA	CTCACT	GTGTTCC	CGGAC	AGACAG	ACAGT	CTGAG	AGGCG	CTTCG	507
QY	253	atgaa	ttccag	catctcc	cag	caat	gtctac	caact	ttg	aaag	312
Db	508	ACA	AGCCCA	AAATG	TGCTA	TGAG	CCAACT	TTG	GGCAC	CTGTCT	567
QY	313	acaag	cccaat	gtgt	ttgt	tg	ccaac	act	tg	gac	372
Db	568	CAAT	ATG	GCCT	CTCTA	TGCA	GAAAC	CTCA	AAABAG	TCTTG	627
QY	373	cag	tata	catc	atct	gtat	gtg	gaac	ctca	aaag	432
Db	628	GAG	CACT	ACTTA	ACCTG	TAAAC	AGAC	ACACT	TAA	GGAAG	687
QY	433	gaac	caact	tta	ccct	gtca	aaag	gaac	atc	gtca	492
Db	688	GCC	AC	CCCT	CTGT	TACT	AGTCC	AAACG	CAAAAG	AGGAT	747
QY	493	gccag	ccct	gt	tact	gt	ctc	ag	gt	tca	552
Db	748	GATT	ATG	CAT	CTGG	GTAT	CAT	TAC	GGCG	TTG	807
QY	553	gat	tc	gc	atc	gtga	tata	ctac	ta	ctg	612
Db	808	AAAA	AA	CACT	TCA	AGG	CACT	TAAT	GAAT	TAT	867
QY	613	aaa	aa	aag	cat	ctca	agga	ca	ataa	ga	672
Db	868	GAC	AA	GAAC	CCCG	CTTAA	AGCT	GCC	AGC	CTCG	927
QY	673	gata	aa	aa	ccg	gcg	caag	ag	ctc	gcag	732
Db	928	ATG	CT	CAAG	TGTG	GAAT	CCAG	AGA	GAAG	GGTGT	987
QY	733	atg	tg	aa	gt	tg	ctcc	gcg	ga	ga	792
Db	988	AGT	T	TAC	CG	AG	CA	GT	AC	T	1047
QY	793	agt	gc	cg	a	gcg	a	gcg	a	gcg	852
Db	1048	CGG	GT	GA	GA	GA	GA	GA	GA	GA	1107
QY	853	cgat	gcg	ga	gcg	gcg	gcg	gcg	gcg	gcg	912
Db	1108	GAA	AG	CT	G	A	C	C	C	C	1167
QY	913	gag	gc	t	gcg	gcg	gcg	gcg	gcg	gcg	972
Db	1168	ATG	A	T	G	A	T	G	A	T	1227
QY	973	atg	a	t	a	t	a	t	a	t	1032
Db	1228	AAG	A	A	A	A	A	A	A	A	1287
QY	1033	aag	a	a	a	a	a	a	a	a	1092
Db	1288	TG	C	T	G	A	T	G	A	T	1347

Tue Nov 3 12:52:16 1998

US-08-906-365-1.rge

Page 17

Db 781 G 781
Qy 1046 g 1046

Search completed: Tue Nov 3 04:15:08 1998
Job time : 2567 secs.

THIS PAGE BLANK (USPTO)

 W I S E R (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPearch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Nov 3 04:51:42 1998; MasPar time 278.07 Seconds

Tabular output not generated. 824.913 Million cell updates/sec

Title: >US-08-906-365-1
 Description: (1-1686) From US08906365.seq
 Perfect Score: 1686
 N.A. Sequence: 1 cagccattatctgcccac.....agaccacagctcagttga 1686
 Comp: gtcggaatattgacgggtg.....tcttggtgctcagagtcac

Scoring table: TABLE default
 Gap 6

Match STD : Dbase 0; Query 0

Searched: 188442 segs, 68026449 bases x 2

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: n-geneseq32
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29 30:part30 31:part31 32:part32 33:part33
 34:part34 35:part35 36:part36 37:part37 38:part38
 39:part39 40:part40

Statistics: Mean 9.451; Variance 5.146; scale 1.837

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	1667	98.9	1936	38	T88415	Human oestrogen recep	0.00e+00
2	1452	86.1	1466	29	T62843	Human oestrogen recep	0.00e+00
3	1434	85.1	1434	37	T88412	Human oestrogen recep	0.00e+00
4	1247	74.0	1251	37	T88413	Human oestrogen recep	0.00e+00
5	1247	74.0	1257	38	T88414	Human oestrogen recep	0.00e+00
6	1136	67.4	2566	29	T62842	Rat oestrogen recep	0.00e+00
7	1010	59.9	1458	29	T62844	Mouse oestrogen recep	0.00e+00
8	197	11.7	4963	17	T06873	PhCMV*-1 promoter and	3.54e-124
9	197	11.7	4963	13	T076270	PhCMV*-1-controlled p	3.54e-124
10	195	7.5	2092	2	N70880	CDNA encoding human o	1.18e-122
11	127	7.5	204	15	Q85976	Human estrogen recept	1.50e-71
12	71	4.2	3460	25	T42193	Fragment of pig oestr	6.97e-31
13	68	4.0	2153	2	N80922	Sequence encoding hum	8.56e-29

14	63	3.7	6244	13	Q76269	PhCMV*-1-controlled p	2.42e-25
15	63	3.7	6244	17	T06872	PhCMV*-1 promoter and	2.42e-25
16	59	3.5	2992	20	T06491	Human hepatoma retino	1.30e-22
17	59	3.5	2992	11	Q65572	Human liver hap CDNA.	1.30e-22
18	59	3.5	2992	13	Q79934	Human liver hap CDNA.	1.30e-22
19	58	3.4	1912	2	Q10388	Encodes skin-specific	6.17e-22
20	58	3.4	2334	38	V05701	Mutant nuclear glucoc	6.17e-22
21	58	3.4	2521	3	Q20542	Mouse retinoic acid r	6.17e-22
22	58	3.4	2740	3	Q20541	Mouse retinoic acid r	6.17e-22
23	58	3.4	5040	2	N80916	Sequence encoding the	6.17e-22
24	56	3.3	1913	18	T05202	Murine retinoic acid	1.38e-20
25	55	3.3	2989	1	N90093	Human Hap (hepatoma)	6.46e-20
26	55	3.3	6176	27	T49226	Plasmid pGR0403r for	1.38e-20
27	56	3.3	7257	31	T68656	Green fluorescent pro	1.38e-20
28	53	3.1	1231	6	Q37761	Sequence of a 1.23 kb	1.40e-18
29	53	3.1	1785	2	Q12005	Human TR2-9 DNA bindi	1.40e-18
30	53	3.1	2029	2	Q12004	Human TR2-5 androgen	1.40e-18
31	53	3.1	2221	2	Q12006	Human TR2-11 DNA bind	1.40e-18
32	53	3.1	2458	2	Q12003	Human TR2-7 DNA bindi	1.40e-18
33	53	3.1	3014	33	T84747	Human progesterone re	1.40e-18
34	53	3.1	3014	35	T94756	Human progesterone re	1.40e-18
35	50	3.0	1351	13	Q79935	Human liver RAR-beta	1.35e-16
36	51	3.0	1376	2	Q10226	Clone pGEM-hRAKgamma	2.96e-17
37	51	3.0	1903	2	Q10389	Encodes human Retinoi	2.96e-17
38	51	3.0	1903	18	T05196	Human retinoic acid r	2.96e-17
39	50	3.0	2249	2	N80921	Sequence encoding hum	1.35e-16
40	51	3.0	3569	29	T63407	Androgen receptor CDN	2.96e-17
41	51	3.0	3569	1	N91772	Human androgen recept	2.96e-17
42	51	3.0	3715	2	Q12001	Full-length human and	2.96e-17
43	49	2.9	204	1	N81164	Base substituted E.co	6.11e-16
44	44	2.9	3762	31	T73334	Human neuron-derived	6.11e-16
45	47	2.8	387	19	T09863	Modified human glucoc	1.23e-14

ALIGNMENTS

RESULT 1
 ID T88415 standard; CDNA; 1936 BP.
 AC T88415:
 DT 20-APR-1998 (first entry)
 DE Human oestrogen receptor protein CDNA #2.
 KW Oestrogen receptor protein; steroid; alternative splicing; estradiol;
 OS Homo sapiens.
 KM estone; estril; screening; ss.
 KW Homo sapiens.
 FH key location/Qualifiers
 FT CDS 77..1669
 FT /tag= a
 FT /product= oestrogen receptor
 FT EP-798378-A2.
 PN 01-OCT-1997.
 PD 25-MAR-1997; 200903.
 PF 22-NOV-1996; EP-203284.
 PR 26-MAR-1996; EP-200820.
 PA (AKU) AKZO NOBEL NV.
 PI DiJkema R, Mosselman S;
 DR WPI: 97-473188/44.
 DR P-PSDB; W33215.
 PT DNA encoding estrogen receptor - useful in screening assay to
 PT identify novel ligands or hormonal analogues
 PS Claim 5; Page 33-35; 45pp; English.
 PS This sequence encodes a novel estrogen binding protein isolated from
 CC human testis CDNA in order to study upstream translation-initiation
 CC codons using 5' RACE-PCR technology. This receptor is able to bind and
 CC be activated by estradiol, estone and estril, can be used in a screening
 CC assay for the identification of new drugs e.g. novel ligands or hormonal
 CC analogues.
 SQ Sequence 1936 BP; 462 A; 514 C; 511 G; 449 T;

Query Match 98.98; Score 1667; DB 38; Length 1936;
 Best Local Similarity 99.98; Pred. No. 0.00e+00;
 Matches 1668; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 cagcatcttggaacattatgaccttctgctctcttctgcaaggtgttcttcta 60

QY	18	caagaa	ctctcttggaacattataatga	cccttctgtgcctctcttcgaa	ggtttctca	77
Db	61	gctgtat	ctccaagacatggataataa	aactccacatagcccttaattc	ctctcttc	120
QY	78	gctctcat	ctccaagacatgataataa	aactccacatagcccttaattc	ctctcttc	137
Db	121	ctaaat	atgagacatcatctta	ccctcttggaacggccatata	caactcttc	180
QY	138	ctaaacat	cgagcaatcatctta	ccctcttggaacggccatata	caactcttc	197
Db	181	ctatgt	agaacagccacatgaa	atalcacagcatctata	gcccctgtgtatga	240
QY	198	ctatgt	agaacagccacatgaa	atalcacagcatctata	gcccctgtgtatga	257
Db	241	ttaaag	atattcccaagcaatg	ctactaacttggaa	gtggcctgtgtgcgagacc	300
QY	258	ttaaag	atattcccaagcaatg	ctactaacttggaa	gtggcctgtgtgcgagacc	317
Db	301	cccaaat	gtgtcttgsgccaacac	ctcttgsgacctctctcctta	gtgtgtcaatgcgcaagt	360
QY	318	cccaaat	gtgtcttgsgccaacac	ctcttgsgacctctctcctta	gtgtgtcaatgcgcaagt	377
Db	361	atcacat	ctgtatgtcggaaac	ctccataaagctccctgtgtga	agcaagatcgtatga	420
QY	378	atcacat	ctgtatgtcggaaac	ctccataaagctccctgtgtga	agcaagatcgtatga	437
Db	421	caactta	cccggtataacgagagac	catctgtgaagaaggtttag	tgggaacccgtctgcgca	480
QY	438	caactta	cccggtataacgagagac	catctgtgaagaaggtttag	tgggaacccgtctgcgca	497
Db	481	ccctgt	tactgtgttcacaggt	ctcaaaaggagatgcacat	ctctgtgcgtctgcgca	540
QY	498	ccctgt	tactgtgtgtccaggt	ctcaaaaggagatgcacat	ctctgtgcgtctgcgca	557
Db	541	cgcact	cggtatctacatgaa	tatgagatctcgtgtgta	agagatgtaagccttttttaag	600
QY	558	cgcact	cggtatctacatgaa	tatgagatctcgtgtgta	agagatgtaagccttttttaag	617
Db	601	aagatc	atccaagagacata	tatataattgtctccaa	gtataaactcagttataac	660
QY	618	aagatc	atccaagagacata	tatataattgtctccaa	gtataaactcagttataac	677
Db	661	aaaccc	ggtgcgaagaagctgt	ccaggtccgtgcggaact	tvggaagtgtacgaagtgtg	720
QY	678	aaaccc	ggtgcgaagaagctgt	ccaggtccgtgcggaact	tvggaagtgtgtgaagtgtg	737
Db	721	gaagt	gtgtgtccctccggag	agagagatgtgtgtga	ccgctctgtgcggaagaa	780
QY	738	gaagt	gtgtgtgtccctccggag	agagagatgtgtgtgtga	ccgctctgtgcggaagaa	797
Db	781	cgacga	gcacagctgacat	ctgtgcggaaggtccaa	gagaaagtgtgcggaac	840
QY	798	cgacga	gcacagctgacat	ctgtgcggaaggtccaa	gagaaagtgtgcggaac	857
Db	841	gcggga	gagctgtgtgtga	acgctctgtgccccga	gacagctaggtccacac	900
QY	858	gcggga	gagctgtgtgtga	acgctctgtgccccga	gacagctaggtccacac	917
Db	901	tgaag	ccgcccccatctgt	ctgtctcaac	tggcccccatccatcat	960
QY	918	tgaag	ccgcccccatctgt	ctgtctcaac	tggcccccatccatcat	977
Db	961	gagt	ctccctgtacgaagt	tggtgcgaagaagattgt	tatacatgtacagctgtgg	1020
QY	978	gagt	ctccctgtacgaagt	tggtgcgaagaagattgt	tatacatgtacagctgtgg	1037
Db	1021	gattcc	cgagctctgtgtga	agcttcagcctgtttc	ggaacgaaggtgcgctctgtg	1080
QY	1038	gattcc	cgagctctgtgtga	agctctcagcctgtttc	ggaacgaaggtgcgctctgtg	1097
Db	1081	gatgtg	aggtgtgtata	tgtgtgtgtgtatgtgtgc	ctaaatttgacaccccgga	1140

Qy	1098	gatgagagtgatgaatgatbaggggctgattgtgcygctcaatttgaccacccccggcaagctcat	
Db	1141	cttggctccagacatctggtctctgacacgggataatgagggaatgocgaaagaaatcttggga	1
Qy	1158	cttggtccccagacatctggtctctgacacgggataatgagggaatgocgaaagaaatcttggga	1
Db	1201	aatcttgacaatgctcctgctgcaactcaacttcaagaagttctcgagaatttaaaatccacaacaa	1
Qy	1218	aatcttgacaatgctcctgctgcaactcaacttcaagaagttctcgagaatttaaaatccacaacaa	1
Db	1261	agaatactctgtgtcaaaagccaatgatcctctgctcaattccaaatgatatgtacctctgtgtac	1
Qy	1278	agaatactctgtgtcaaaagccaatgatcctctgctcaattccaaatgatatgtacctctgtgtac	1
Db	1321	agcaaccacgaatgctgctgacacagcccggaagcttggtctcaactgtgctgcaacgcgtgaccca	1
Qy	1338	agcaaccacgaatgctgctgacacagcccggaagcttggtctcaactgtgctgcaacgcgtgaccca	1
Db	1381	tgcttggttgttggatgtccaaagacggcatalctctccacagcaatccatgagctct	1
Qy	1398	tgcttggttgttggatgtccaaagacggcatalctctccacagcaatccatgagctct	1
Db	1441	ggctaacctctctgatgctcctctgctccacgtcgaatgagtgagtaacaaaggcattgaaaca	1
Qy	1458	ggctaacctctctgatgctcctctgctccacgtcgaatgagtgagtaacaaaggcattgaaaca	1
Db	1501	tctgtcaacatgaatgctgacaaatgtgtgtccagtgatactgacccctgcctgcctggaatgct	1
Qy	1518	tctgtcaacatgaatgctgacaaatgtgtgtccagtgatactgacccctgcctgcctggaatgct	1
Db	1561	gaatgcccacgtgctctcgcggtgtgcgaagctctccatcaacaggggtccgagtgacgccggc	1
Qy	1578	gaatgcccacgtgctctcgcggtgtgcgaagctctccatcaacaggggtccgagtgacgccggc	1
Db	1621	agaaggaacgtacaagaagcgaagggctcccaagaaccacagctctcaagta	1669
Qy	1638	agaaggaacgtacaagaagcgaagggctcccaagaaccacagctctcaagta	1686


```
433 gaacacacttaacctgtataacagagagacactgaagaaggaaggttaagcggagacggttgc 492
Db 433 gcaacacacttaacctgtataacagagagacactgaagaaggaaggttaagcggagacggttgc 492
Qy 241 gccagccctgtactgtgtccaggtgtcaaaagagagatgtctcaactctgtcgtctgcagc 300
Qy 493 gccagccctgtactgtgtccaggtgtcaaaagagagatgtctcaactctgtcgtctgcagc 552
Db 301 gattaacgcatcgagatatacactataagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 360
Qy 553 gattacgcatcgagatatacactataagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 612
Db 361 aaaagaacatcataagagacataatgattatttttccagctacataacacagtgtaacac 420
Qy 613 aaaagaacatcataagagacataatgattatttttccagctacataacacagtgtaacac 672
Db 421 gataaaaacgagcgcaaaagatgtccagagcgctgcgactctcgaaagtgttacaagttgga 480
Qy 673 gataaaaacgagcgcaaaagatgtccagagcgctgcgactctcgaaagtgttacaagttgga 732
Db 481 atggtgaagtgtgtctcccgagagagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 540
Qy 733 atggtgaagtgtgtctcccgagagagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 792
Db 541 agtgcgagcgagcgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 600
Qy 793 agtgcgagcgagcgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 852
Db 601 cgaatgcgagcgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 660
Qy 853 cgaatgcgagcgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 912
Db 661 gagagctgaagcgcccgacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 720
Qy 913 gagagctgaagcgcccgacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 972
Db 721 atggtgaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 780
Qy 973 atggtgaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1032
Db 781 aagaagatctccgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 840
Qy 1033 aagaagatctccgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1092
Db 841 tgtgtgaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 900
Qy 1093 tgtgtgaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1152
Db 901 ctcatcttctgtccagatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
Qy 1153 ctcatcttctgtccagatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1212
Db 961 ctggaatctctgtacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1020
Qy 1213 ctggaatctctgtacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1272
Db 1021 cacaagaatatactctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1080
Qy 1273 cacaagaatatactctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1332
Db 1081 gtcaacagcgacacagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1140
Qy 1333 gtcaacagcgacacagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1392
Db 1141 accgatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1200
Qy 1393 accgatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1452
Db 1201 cgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1247
Qy 1453 cgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1499
```

```
ID T88414 standard; cDNA; 1257 BP.
AC T88414;
DT 20-APR-1998 (first entry)
DE Human oestrogen receptor protein splice variant cDNA (exon 8c).
KW Oestrogen receptor protein; steroid; alternative splicing; estradiol;
KW estrone; estriol; screening; ds.
OS Homo sapiens.
PN EP-798378-A2.
PD 01-OCT-1997.
PF 25-MAR-1997; 200903.
PR 22-NOV-1996; EP-203284.
PR 26-MAR-1996; EP-200820.
PA (ALKU ) AKZO NOBEL NV.
PI Dijkema R, Mosselman S;
DR WPI: 97-473188/44.
PT P-PSDB: W33214.
PR DNA encoding estrogen receptor - useful in screening assay to
PT identify novel ligands or hormonal analogues
PS Claim 5; Page 28-29; 45pp; English.
CC This sequence encodes splice variant of a novel oestrogen binding protein
CC isolated from human thymus tissue. This protein contain an alternative
CC exon 8 (exon 8c) of the novel oestrogen receptor represented in T88412.
CC This novel receptor is able to bind and be activated by estradiol, estrone
CC and estriol, can be used in a screening assay for the identification of
CC new drugs e.g. novel ligands or hormonal analogues. This variant does not
CC contain an AF-2 region and therefore probably lacks the ability to
CC modulate transcription of target genes in a ligand dependent fashion.
SQ Sequence 1257 BP; 297 A; 327 C; 355 G; 278 T;

Query Match 74.0%; Score 1247; DB 38; Length 1257;
Best local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 1247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 atgattacacatctccagcaatgtctactactgtgaagttggtgctgtgcagacc 60
Qy 253 atgattacacatctccagcaatgtctactactgtgaagttggtgctgtgcagacc 312
Db 61 acaagcccaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 120
Qy 313 acaagcccaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 372
Db 121 caattatcacatctgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 180
Qy 373 caattatcacatctgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 432
Db 181 gaacacacatctactgttaaacagagagacactgtgaagaagttaagtgtgaacgtttgc 240
Qy 433 gaacacacatctactgttaaacagagagacactgtgaagaagttaagtgtgaacgtttgc 492
Db 241 gccagccctgtactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 300
Qy 493 gccagccctgtactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 552
Db 301 gattaacgcatcgagatatacactataagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 360
Qy 553 gattaacgcatcgagatatacactataagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 612
Db 361 aaaagaacatcataagagacataatgattatttttccagctacataacacagtgtaacac 420
Qy 613 aaaagaacatcataagagacataatgattatttttccagctacataacacagtgtaacac 672
Db 421 gataaaaacgagcgcaaaagatgtccagagcgctgcgactctcgaaagtgttacaagttgga 480
Qy 673 gataaaaacgagcgcaaaagatgtccagagcgctgcgactctcgaaagtgttacaagttgga 732
Db 481 atggtgaagtgtgtctcccgagagagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 540
Qy 733 atggtgaagtgtgtctcccgagagagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 792
Db 541 agtgcgagcgagcgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 600
Qy 793 agtgcgagcgagcgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 852
```

DB 601 cgaatgcggagacgtctgtctgacgcgccttgagcccccagacagtagtctacccctctg 660
|||||
OY 853 cgaatgcggagacgtctgtctgacgcgccttgagcccccagacagtagtctacccctctg 912
DB 661 gagatgcggagacgtctgtctgacgcgccttgagcccccagacagtagtctacccctctg 720
913 gagatgcggagacgtctgtctgacgcgccttgagcccccagacagtagtctacccctctg 972
DB 721 atgagatgcgccttgacgcgccttgagcccccagacagtagtctacccctctg 780
OY 973 atgagatgcgccttgacgcgccttgagcccccagacagtagtctacccctctg 1032
DB 781 aagaagatcccggtctgttgagcctcagcctgttcgacaaagtcgagcctcttgagagc 840
OY 1033 aagaagatcccggtctgttgagcctcagcctgttcgacaaagtcgagcctcttgagagc 1092
DB 841 ttttgatgagagtggttaataatgagtggtgagtggtgagtggtgagtggtgagtggtgag 900
OY 1093 ttttgatgagagtggttaataatgagtggtgagtggtgagtggtgagtggtgagtggtgag 1152
DB 901 cccatcttgccacagatctgtctgacgcgccttgagcccccagacagtagtctacccctctg 960
OY 1153 cccatcttgccacagatctgtctgacgcgccttgagcccccagacagtagtctacccctctg 1212
DB 961 ctggaatctctgacatgctcctgacacacttcaaggtttcgaagatlaaactcaaa 1020
OY 1213 ctggaatctctgacatgctcctgacacacttcaaggtttcgaagatlaaactcaaa 1272
DB 1021 ccaaaagaatattctgtctgacgcgccttgagcccccagacagtagtctacccctctg 1080
OY 1273 ccaaaagaatattctgtctgacgcgccttgagcccccagacagtagtctacccctctg 1332
DB 1081 gtcacagcagccagatctgacacagcagcagcagcagcagcagcagcagcagcagcagc 1140
OY 1333 gtcacagcagccagatctgacacagcagcagcagcagcagcagcagcagcagcagcagc 1392
DB 1141 accgagatctgt 1200
OY 1393 accgagatctgt 1452
DB 1201 cgcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1247
OY 1453 cgcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1499

RESULT 5
ID T62842 standard; cDNA: 2568 BP.
AC T62842:
DE 08-JUN-1997 (first entry)
DE Rat oestrogen receptor beta (ER-beta) cDNA.
KW Orphan receptor; oestrogen receptor beta; ER-beta;
KW Orphan receptor; prostate cancer; benign prostatic hyperplasia;
KW ovarian cancer; cardiovascular disease; osteoporosis;
KW environment; pollutant; ss.
OS Rattus sp.
FH Key Location/Qualifiers
FT cds 424..1881
FT /tag= a
FT /note= "alternative start codon at 448..450"

MOJ09348-A2.
PD 13-MAR-1997.
PE 09-SEP-1996; E03933.
PR 08-SEP-1995; GB-018272.
PR 15-MAR-1996; GB-005550.
PR 11-APR-1996; GB-007532.
PR 08-MAY-1996; GB-009576.
PA (KARO-) KARO BIO AB.
PI Enmark E, Gustafsson J, Kuiper GG;
DR WPI: 97-192842/17.
DR P-PDB: W14723.
PT New isolated oestrogen receptor beta - used to develop prods. for
PT treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT disease

PS Claim 6; Fig 1; 45pp; English.
CC A cDNA clone (T62842) codes for a novel rat oestrogen receptor-
CC related nuclear receptor (W13723), designated ER-beta. A set of
CC degenerate primers based on conserved sequences of the DNA-
CC binding domain and ligand binding domain of nuclear receptors were
CC used to amplify rat prostate cDNA. The PCR prod. was used to
CC screen a rat prostate cDNA library to obtain the ER-beta cDNA.
CC This was then used to obtain human ER-beta cDNA (see also T62843)
CC from a human testis cDNA library. Rat, human and mouse ER-beta
CC (W14723-25) can be used to isolate molecules for use in the
CC treatment of cardiovascular diseases, central nervous system
CC diseases, osteoporosis, prostate or ovarian cancer or benign
CC prostatic hyperplasia and to test environmental chemicals for
CC oestrogenic activity.
SQ Sequence 2568 BP; 634 A; 654 C; 663 G; 617 T;

Query Match 67.4%; Score 1136; DB 29; Length 2568;
Best Local Similarity 84.8%; Pred. No. 0.00e+00;
Matches 1420; Conservative 0; Mismatches 248; Indels 6; Gaps 5;

DB 211 cccatgagctctctgagacataatg-tc-catctgacctctctcacaagagtttct 268
OY 16 cccagagatctcttgagacataatgacatcttgctctctcttgcaagtggtttct 75
DB 269 cagctgcagccctctgagacatgagatcaaaaactcaccgctgcagccttagtc-ctt 327
OY 76 cagctgcatac-tc--aagacatgagataaaaactcaccatctagccttaattctct 132
DB 328 gcttcttaataatgtagccagatccatctctacccctggagcagcccatcattcct 387
OY 133 tctctctaaatctgagatcaatccatctctacccctggagcagcagcagcagcagcagc 192
DB 388 tctctctaaatctgagac 447
OY 193 tctctctaaatctgagac 252
DB 448 atgaaactagtggtccgcgcagcagcagcagcagcagcagcagcagcagcagcagcagc 507
OY 253 atgaaactagtggtccgcgcagcagcagcagcagcagcagcagcagcagcagcagcagc 312
DB 508 acaagcccaaatggtctatgagcagcagcagcagcagcagcagcagcagcagcagcagc 567
OY 313 acaagcccaaatggtctatgagcagcagcagcagcagcagcagcagcagcagcagcagc 372
DB 568 caatcagctctctctatgagcagcagcagcagcagcagcagcagcagcagcagcagc 627
OY 373 cagttatcagatctgtatgagcagcagcagcagcagcagcagcagcagcagcagcagc 432
DB 628 gacacacacttaactgttaaacagagacacacacacacacacacacacacacacacac 687
OY 433 gacacacacttaactgttaaacagagacacacacacacacacacacacacacacacac 492
DB 688 gccagccctgttaccagttccaaagcagagagagtgctcatttgcgcgtctgcagc 747
OY 493 gccagccctgttaccagttccaaagcagagagagtgctcatttgcgcgtctgcagc 552
DB 748 gattatgacatggttatcattacacgagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 807
OY 553 gattatgacatggttatcattacacgagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 612
DB 808 aaaagaagcattcaagagacataatgattatctctgcccagcagcagcagcagcagcagc 867
OY 613 aaaagaagcattcaagagacataatgattatctctgcccagcagcagcagcagcagcagc 672
DB 868 gacaaagacagcgtcaaaagcgtccagcagcgtccagcagcgtccagcagcgtccagcagc 927
OY 673 gataaaacacggtcgaagcagcgtccagcagcgtccagcagcgtccagcagcgtccagcagc 732
DB 928 atggtcaagtggtatccagcagcagcagcagcagcagcagcagcagcagcagcagcagc 987
OY 733 atggtcaagtggtatccagcagcagcagcagcagcagcagcagcagcagcagcagcagc 792
DB 988 agttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1047

[illegible]

RESULT	7
ID	T62844 standard; cDNA; 1458 BP.
AC	T62844;
DT	08-JUN-1997 (first entry)
DE	Mouse oestrogen receptor beta (ER-beta) cDNA.
KW	Orphan receptor; oestrogen receptor beta; ER-beta;
KW	nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW	ovary cancer; cardiovascular disease; osteoporosis;
KW	environment; pollutant; ss.
OS	Mus sp.
PN	W09705348-A2.
PD	13-MAR-1997.
PF	09-SEP-1996; E03933.

PR 08-SEP-1995; GB-018272.
PR 15-MAR-1996; GB-005550.
PR 11-APR-1996; GB-007532.
PR 08-MAY-1996; GB-009576.
PA (KARO-) KARO BIO AB.
PI Enmark E, Gustafsson J, Kuiper GG;
DR WPI; 97-192842/17.
P-PSDB; W14725.
PT New isolated oestrogen receptor beta - used to develop prods. for
PT treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT disease
PS Claim 6; Fig 14B; 45pp; English.
CC A cDNA clone (Tc2844) codes for a novel mouse oestrogen receptor-
CC related nuclear receptor (W13725), designated ER-beta. Rat, human
CC and mouse ER-beta (W14723-25) can be used to isolate molecules for
CC use in the treatment of cardiovascular diseases, central nervous
CC system diseases, osteoporosis, prostate or ovarian cancer or benign
CC prostatic hyperplasia and to test environmental chemicals for
CC oestrogenic activity.
SQ Sequence 1458 BP; 365 A; 373 C; 403 G; 317 T;

Query Match	59.9%	Score 1010	DB 29	Length 1458
Best Local Similarity	84.6%	Pred. No. 0.00e+00		
Matches 1234	Conservative	0	Mismatches 224	Indels 0
			Gaps	
Db	1	atggaattctcaagcctcgtctgtatgatactacagatgtgtccagcagcacccggtacaactg	60	
QY	229	atgaattctctatagccctcgtctgtatgatactacagcatctccagcaatgttactaacttg	288	
Db	61	gaaggttgccgtgtctgcgcgaagcttgcgaagcccaatgtgtctatgtgcgaactctgcgac	120	
QY	289	gaaggttgccgtgtcgcgcagaccacaagcccaatgtgtgtgtgccaacccctgycac	348	
Db	121	ctctctctcttagccaccacattgcacatcgtctctctatcgcagaactccataaagat	180	
QY	349	cttctctctcttagtgtctcaccgcagttatctacatctgtatgtgycgaactccataaagat	408	
Db	181	cctgtgtgtgaagcaagaatcactagaacaacacctgtgcctgttaaacagagagacctgaag	240	
QY	409	ccctgtgtgtgaagcaagatgtctctagaacacaccttacctgttaaacagagagacctgaag	468	
Db	241	aggaagcttgcgcgcgcgcgcgtgtgtgcgcgcctgttctactagttcccaagcaccagaagagat	300	
QY	469	aggaaggttagtgggaaacgcttgcgcgcgcgcctgttactgtgtccagttccaagagagat	528	
Db	301	gtcaactctctgtccgctctgtcgaatgtatgtcatctctgtgtatcatctacgtgtgtgtcc	360	
QY	529	gtcaactctctgtccgctctgtcgaatgtatgtcatctctgtgtatcatctacgtgtgtgtcc	588	
Db	361	tgtgaagatgttaagccctttttaaagaagcatttaagagacataatgatactgt	420	
QY	569	tgtgaagatgttaagccctttttaaagaagcatttaagagacataatgatactgt	648	
Db	421	ccagcccaagaaatcagctgtctagatagacaagaaccgcgttaaaaaacttccagccttcgca	480	
QY	649	ccagctacaaatcagctgtctagatagacaagaaccgcgttaaaaaacttccagccttcgca	708	
Db	481	cttcgcgaagtgttaccgaaatgatagtgatacgtgtggatccagagagaaagtgtggg	540	
QY	709	cttcgcgaagtgttaccgaaatgatagtgatacgtgtggatcccgagagaaagtgtggg	768	
Db	541	taccgaatagatagcaagaagagagagatgtccgcgcgaagcagctgtccactgtgccggaaggcc	600	
QY	769	taccgcctgtgtgcgcgcgaagagagagatgtccgcgcgaagcagctgtccactgtgccggaaggcc	828	
Db	601	aagaaacacagttgggcacaacaccccggtgtgaagagatctatgtctgaactctctgaagcc	660	
QY	829	aagaaagttgggcgcgcacgc	888	
Db	661	gagcagcctgtgtgtccacccctctgtgaaagctgcagccaccacaatgtgtcagtgtgtctcc	720	
QY	889	gagcagcctgtgtgtccacccctctgtgaaagctgcagccaccacaatgtgtcagtgtgtctcc	948	

WIDEORH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Nov 2 16:27:57 1998; MasPar time 27.56 Seconds

Tabular output not generated. 702,514 Million cell updates/sec

Title: >US-08-906-365-2

Description: (1-530) from US08906365.pep

Perfect Score: 3904

Sequence: 1 MDIKNSPSSINSPPSYNCSQ.....ECSPAEDSKSEGSQNPQSQ 530

Scoring table: PAM 150

Gap 11

Searched: 120441 seqs, 3653193 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: p1r56

1:p1r1 2:p1r2 3:p1r3 4:p1r4 5:nr13d

Statistics: Mean 48.751; Variance 99.595; scale 0.489

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	3532	90.5	477	2	S71400 estrogen receptor bet	0.00e+00
2	1735	44.4	586	1	ORXLE estrogen receptor - A	0.00e+00
3	1710	43.8	589	1	ORCHE estrogen receptor - C	0.00e+00
4	1683	43.1	595	1	ORHUE estrogen receptor 1 -	0.00e+00
5	1669	42.8	599	1	ORHSE estrogen receptor - m	0.00e+00
6	1663	42.6	600	1	ORRTE estrogen receptor - r	0.00e+00
7	1654	42.4	595	2	I47140 estradiol receptor - r	5.58e-303
8	1516	38.8	574	2	A37197 estrogen receptor - r	1.66e-274
9	1503	38.5	535	2	S58224 oestrogen receptor - r	7.90e-272
10	1473	37.7	701	2	S64737 80K estrogen receptor	1.15e-265
11	807	20.7	433	2	B29345 steroid hormone recept	4.93e-130
12	799	20.5	433	2	S58087 estrogen receptor rel	1.97e-128
13	800	20.5	521	2	A29345 steroid hormone recep	1.24e-128
14	681	17.4	923	2	A39596 progesterone receptor	6.28e-105
15	671	17.2	923	2	I53280 progesterone receptor	5.95e-103
16	666	17.1	933	1	ORHUP progesterone receptor	5.84e-102
17	662	17.0	930	2	A25923 progesterone receptor	3.61e-101
18	614	15.7	742	1	ORHUGA glucocorticoid recept	1.04e-91
19	614	15.7	777	1	ORHUGA glucocorticoid recept	1.04e-91
20	608	15.6	776	1	S44047 glucocorticoid recept	1.57e-80
21	597	15.3	786	2	A35466 progesterone receptor	2.25e-88
22	594	15.2	771	2	A54273 glucocorticoid recept	8.72e-88
23	586	15.0	783	1	A25691 glucocorticoid recept	3.21e-86

24	579	14.8	133	2	S26595 estrogen receptor - b	7.52e-85
25	577	14.8	795	1	ORRTG glucocorticoid recept	1.85e-84
26	571	14.6	899	2	A35895 androgen receptor - m	2.75e-83
27	570	14.6	902	2	B40494 androgen receptor - r	4.31e-83
28	567	14.5	919	2	A39248 androgen receptor - h	1.66e-82
29	564	14.4	910	2	A34721 androgen receptor A -	6.39e-82
30	562	14.4	981	2	A41401 mineralocorticoid rec	1.57e-81
31	557	14.3	74	5	1HCOA humanCHICKEN ESTROGEN	1.48e-80
32	557	14.3	74	5	1HCOE humanCHICKEN ESTROGEN	1.48e-80
33	557	14.3	75	5	1HCP humanCHICKEN ESTROGEN	1.48e-80
34	560	14.3	911	2	B34721 androgen receptor B -	3.85e-81
35	547	14.0	71	5	1HCOB humanCHICKEN ESTROGEN	1.31e-78
36	543	13.9	344	2	I51330 androgen receptor - c	7.84e-78
37	544	13.9	758	2	S60586 glucocorticoid recept	5.01e-78
38	533	13.7	533	2	S37781 retinoid x receptor b	6.86e-76
39	529	13.6	410	2	S26669 retinoic acid recepto	4.10e-75
40	529	13.6	446	2	A34418 H-2 region II binding	4.10e-75
41	529	13.6	448	2	D41727 retinoid x receptor b	4.10e-75
42	529	13.6	451	2	A41651 retinoic acid recepto	4.10e-75
43	523	13.4	984	2	A29513 mineralocorticoid rec	5.96e-74
44	514	13.2	467	2	S26668 retinoic acid recepto	3.29e-72
45	514	13.2	467	2	A47278 retinoid x receptor a	3.29e-72

ALIGNMENTS

RESULT 1
ENTRY 1 S71400 #type complete
TITLE estrogen receptor beta - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 26-Feb-1998

ACCESSIONS S71400
REFERENCE S71400
#authors Mosselman, S.; Polman, J.; Dijkema, R.
#journal FEBS Lett. (1996) 392:49-53
#title ER-beta: identification and characterization of a novel human estrogen receptor.

#accession S71400

#molecule_type mRNA
#residues 1-477 #label MOS

#cross-references EMBL:X99101

CLASSIFICATION #superfamily estrogen receptor; erba transforming protein homology

KEYWORDS DNA binding; nucleus; phosphoprotein; steroid binding; steroid hormone receptor; transcription regulation; zinc finger

FEATURE 94-355 #domain erba transforming protein homology #label ERBA\

96-116 #region zinc finger CCCC motif\

132-156 #region zinc finger CCCC motif\

167-182 #region nuclear location signal\

52 #binding_site phosphate (Ser) (covalent) (by MAP kinase)

96,99,113,116 #binding_site zinc (Cys) #status predicted\

132,138,148,151 #binding_site zinc (Cys) #status predicted\

SUMMARY #length 477 #molecular_weight 53363 #checksum 8603

Query Match 90.5%; Score 3532; DB 2; Length 477;

Best local similarity 100.0%; Pred. No. 0.00e+00;

Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB	1	MNYSIPSVNTNLEGGPGQTSPNVLMPTPGHLSLVVHRQLSHYAPQSPWCEASL	60
OY	54	MNYSIPSVNTNLEGGPGQTSPNVLMPTPGHLSLVVHRQLSHYAPQSPWCEASL	113
DB	61	EHLLPVNRETLKRRKSGNRCASPVTPGSPKRDARHCACVDSAGYHGWSCCEGKAF	120
OY	114	EHLLPVNRETLKRRKSGNRCASPVTPGSPKRDARHCACVDSAGYHGWSCCEGKAF	173
DB	121	KRSIOGHNDYICPATNOCTIDKNRRKSCQACRLKRCYEVGWKCGSRRERCGYRLVRQR	180
OY	174	KRSIOGHNDYICPATNOCTIDKNRRKSCQACRLKRCYEVGWKCGSRRERCGYRLVRQR	233

Db	181	SAEOLHCAGAKKSGGAPRVRELLDALSPEOLVTLLEAEPPHILSRPAPFTKMS	240
Qy	234	SADBDLHCAGAKKSGGAPRVRELLDALSPEOLVTLLEAEPPHILSRPAPFTKMS	293
Db	241	MMMSLTKLADKELYHMISMAKKIPGEVBLSLFDQVRLLESCWMEVLMGMIMKMSIDHPGK	300
Qy	294	MMMSLTKLADKELYHMISMAKKIPGEVBLSLFDQVRLLESCWMEVLMGMIMKMSIDHPGK	353
Db	301	LIFAPDVLDRDEGCVEGILIEIFDMLLATTSRRELKLOHKKEYLCVKAMILLNSSMYPL	360
Qy	354	LIFAPDVLDRDEGCVEGILIEIFDMLLATTSRRELKLOHKKEYLCVKAMILLNSSMYPL	413
Db	361	VTATODADSSKRLAHLNAVTDALVWYIAKGGISSQOOSMRALMLMLLSHYVHNASKGM	420
Qy	414	VTATODADSSKRLAHLNAVTDALVWYIAKGGISSQOOSMRALMLMLLSHYVHNASKGM	473
Db	421	EHLNMKKNNVPYVDLLEMLNHAIVLGCKSSITGSECSAPBEDSKSEGSQNPQSQ	477
Qy	474	EHLNMKKNNVPYVDLLEMLNHAIVLGCKSSITGSECSAPBEDSKSEGSQNPQSQ	530

RESULT	2
ENTRY	ORXLE
TITLE	#type complete
ORGANISM	estrogen receptor - African clawed frog
DATE	#formal name Xenopus laevis #common name African clawed frog
	28-Feb-1992 #sequence_revision 14-Jul-1994 #text_change 15-Jun-1996
ACCESSIONS	A40907
REFERENCE	A40907
#authors	Weiler, I.J.; Lew, D.; Shapiro, D.J.
#journal	Mol. Endocrinol. (1987) 1:355-362.
#title	The Xenopus laevis estrogen receptor: sequence homology with

#cross-references MUID:90331927

```
##molecule_type mrna
##residues 1-586 ##label WEI
##cross-references GB:L20735
```

COMMENT The steroid hormones and their receptors are involved in the regulation of eukaryotic gene expression and affect cellular proliferation and differentiation in target tissues.

COMMENT In the absence of ligand, steroid hormone receptors are thought to be weakly associated with nuclear components; hormone binding greatly increases receptor affinity. The hormone-receptor complex appears to recognize target DNA sequences upstream of transcriptional start sites.

CLASSIFICATION #superfamily estrogen receptor; erba transforming protein homology

KEYWORDS DNA binding; nucleus; phosphoprotein; steroid binding; steroid hormone receptor; transcription regulation; zinc finger

	FEATURE
1-174	#domain amino-terminal #label NH2\
175-266	#domain DNA binding #status predicted #label DNA\
178-448	#domain eba transforming protein homology #label EBBA\
178-201	#region zinc finger CCCC motif\
214-237	#region zinc finger CCCC motif\
251-266	#region nuclear location signal\
293-554	#domain steroid binding #status predicted #label STR\
180,183,197,200	#binding site zinc (Cys) #status predicted\
216,222,232,235	#binding site zinc (Cys) #status predicted\
231	#binding-site phosphate (Ser) (covalent) #status predicted\

SUMMARY	#Length	586	#molecular-weight	66080	#checksum	9991
Query Match		44.4%	Score	1735;	DB 1;	Length 586;

Best Local Similarity 50.8%; Pred. No. 0.00e+00;
Matches 247; Conservative 112; Mismatches 107; Indels 20; Gaps 16

```

Db      70 YSSASLSYASSETFG-SSSLTGLHT-LNNVPPSPVYELAKLPOLSPFIHHGQQVPEYYL 127
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

OY	43	YPAATF--YSPAVNMYSPISVNTJNEGPRQOTSPNVAJMTPEBLSLYVNR--OLSHLYA	100
Db	128	ESBEGT-AVREARPFIFYRSSDNRRSGREHMSANDKPPSMSTKRETCYAVOSDY	186
OY	101	EPQSPMCEARSLHEJLPVNRBEJTKRVSG--NR-CASPVTGP-G--SKRAHCAVOSDY	155
Db	187	ASGHYGVWSCGCGKAFPKRSIOGHNDYMCPATNOCTIDKNRRKSCOACRLKCYEYGM	246
OY	156	ASGHYGVWSECGCKAFPKRSIOGHNDYICPATNOCTIDKNRRKSCOACRLKCYEYGMV	215
Db	247	KGGIRKDRGGRLLKHNRKOEBOKRDVDPSELRTSJMWNVYSKMKSLPYLSIAEQ	306
OY	216	KCGSRRECCGRILRV--RQSSADQLH-CAGKAKRSG--HAPRYRLILDA-LS--PEQ	267
Db	307	LISALMEAEAVIYSEHSDTKPESLEASAMTLLTNADRELVHIMINAKRVPGFVDLTHD	366
OY	268	LVLTLLEAPPHVILISRPSA--PFEASMMMSITKLADRELVHIMISMAKKTIPGEVELSLFD	326
Db	367	QVHLLECAMLEILWGLIWSVEHPGLSPFNILLDRNGRCYEGJVELFEDMIJYTAATR	426
OY	327	QVRLLSECMVMVLMMGLMWRSIDHPGLIAPDVLDRDESKCYEGLLEFDMILATITS	386
Db	427	FRMRRLRGEETICLSIILLNSGYUTFLSTJESLEDTDLHLILDKIIDPVLVFNMAKSG	486
OY	387	FRELKTOIKETLCYKAMILLNMSMYPLVTANQDADSSKRLAHL--LNVDTALVWVIKSG	445
Db	487	LSLOOQORRLAOLLILSHITFMSNKGMEHLISYKCKKNVPLVJLLLEMDIAHITPKD	546
OY	446	ISSQOOSRRLANLMLLSHVHNASNKGMEHLINCKKNVAVPYDLLLEMANAHVLRCKS	505
Db	547	KTTTQE 552	
OY	506	SITGSE 511	

RESULT	3	QRCHE	#type complete
ENTRY		estrogen receptor - chicken	
TITLE		#formal_name Gallus gallus	#common_name chicken
ORGANISM		06-Mar-1992	#sequence_revision 14-Jul-1994
DATE		06-Mar-1992	#text_change

ACCESSIONS REFERENCE	05-Sep-1997
A40914; S071922	
A40914	

#authors Maxwell, G.L.; McDonnell, D.L.; Conner, L.S.; Greene, G.L.; O'Valley, B.W.
#journal Mol. Endocrinol. (1987) 1:25-35
#title Structural organization and regulation of the chicken estrogen receptor.
#cross-references MUID:88318621

#accession	A40914	
##status	not compared with conceptual translation	
#molecule_type	mRNA	
##residues	1-589	##label MAX

REFERENCE	#authors
50/192	
Krust, A.; Green, S.; Argos, P.; Kumar, V.; Walter, P.; Bornert, J.M.; Chandon, P.	
EMBO J. (1996) 5, 801-807	

#journal EMBO J. (1997) 16:351-62
#title The chicken oestrogen receptor sequence: homology with v-erbA
and the human oestrogen and glucocorticoid receptors.
#cross-references MUID:86247578

```

#accession      S07192
##status        preliminary
##molecule-type  mRNA
##residues      1-256, 'E', 258-589 ##label  KRU
##cross-references EMBL:X03805; NID:G63378; PID:G63380
#Comments
#steroid hormone and their receptors are involved in

```

COMMENT: Steroid hormones and their receptors are involved in the regulation of eukaryotic gene expression and affect cellular proliferation and differentiation in target tissues.

COMMENT: In the absence of ligand, steroid hormone receptors are thought to be inactive and to be sequestered in the cytoplasm by heat shock proteins.

the weakly associated with nuclear components; hormone binding greatly increases receptor affinity. The hormone-receptor complex appears to recognize discrete DNA sequences upstream of transcriptional start sites.

CLASSIFICATION #superfamily estrogen receptor; erba transforming protein
 KEYWORDS DNA binding; nucleus; phosphoprotein; steroid binding; steroid hormone receptor; transcription regulation; zinc finger

FEATURE
 1-173 #domain amino-terminal #label NH2
 174-255 #domain DNA binding #status predicted #label DNA
 177-450 #domain erba transforming protein homology #label ERBA
 213-235 #region zinc finger CCCC motif
 230-265 #region zinc finger CCCC motif
 250-265 #region nuclear location signal
 294-566 #domain steroid binding #status predicted #label STB
 179,182,196,199 #binding_site zinc (Cys) #status predicted
 215,221,231,234 #binding_site zinc (Cys) #status predicted
 230,299 #binding_site phosphate (Ser) (covalent) #status predicted

SUMMARY #length 589 #molecular-weight 66773 #checksum 793

Query Match 43.8%; Score 1710; DB 1; Length 589;
 Best Local Similarity 51.6%; Pred. No. 0.00e+00;
 Matches 247; Conservative 112; Mismatches 99; Indels 21; Gaps 14;

Db 69 YGSLTSLVPTSESEF-SSSLAGFHS-LNNVPPSPVYFLOTAPQLSPFHHSQVPIYL 126
 43 YPAMTF-YSPAVMNYSPISPNVNTLEGGPGRQTTSPNVLPTEHLSPLVYHR-QLSHLYA 100
 127 ENEQSGFGRAPPAFPYRSPSSDNRRHSIRERNSSFNKESLSMESTKTRCAVANDYA 186
 101 EPQKSPWCARSLERLT-PVNNETLKRKVSNGNCASPVTPG-G-SKRDHFCFVACSDYA 156
 187 SCYHYGVMSCEGCKAFKFKSIQGHNDYMCPTNOCTIDKNNRKSQACRLKCYEVMK 246
 157 SCYHYGVMSCEGCKAFKFKSIQGHNDYICPATNOCTIDKNNRKSQACRLKCYEVMK 216
 247 GGRDRGRGRMKQRQREEDSNGEASSTELRAPLMTSPLELVYKHNKNSPALSLTA 306
 217 CSRRRCRCYLRVRRORSADEQLHCAGKAKRSQGHAPRV-R-E-LTL-----D--ALS--P 265
 307 EOMVSLLEAPPIYSEVDPNRPNENASMTLLTNLADRELHYHMLNNAKRVGFYDILT 366
 266 EDVLTLLEAPPPHVLIS-RPSAPTEASMMSLTKLADRELHMLSMKKKIPGFELSL 324
 367 HDQVHLECAMLEIIMTGLVMSMEHPKGLFAPNLLDRNKGCEVGEVEIFDMLATA 426
 325 PQVRLTSCMVEVLMGMWRSIDHPKGLIFAPDLVDRDEKCVCEGILEIFDMLATT 384
 427 AFRPMNLTGEEFVCLSIILNSGYTFLSTLSLEERDYIHRVLDKITDTLILHMAK 486
 385 SFRRELQKHKEYLCKVAMILLNSSMYPLVTATQDADSSRKLAH-LLNATVDALVWVIAR 443
 487 SGLSLQOQRRLAOLLLIISHIRHMSKGMHEHLYNMCKKNVYVLYLLEMDAHLHA 545
 444 SGISSQOQMRLANLMLLSHVHSHASNKGMHEHLNCKKNVYVYDLEMLNAHVL 502

RESULT 4
 ENTRY 4
 TITLE ORHUE #type complete
 ALTERNATE_NAMES estrogen receptor 1 - human
 ORGANISM ERI; estrogen receptor alpha
 DATE #journal, name Homo sapiens #common_name man
 28-May-1986 #sequence_revision 28-May-1986 #text_change
 20-Mar-1998
 ACCESSIONS A94284; A93376; A43021; S27143; S34000; A41925; B41925;
 A03244; C41925; D41925
 REFERENCE A94284
 #authors Greene, G.L.; Gilna, P.; Waterfield, M.; Baker, A.; Hort, Y.;
 #journal Shine, J.
 Science (1986) 231:1150-1154
 #title Sequence and expression of human estrogen receptor
 complementary DNA.
 #cross-references MUID:86122927
 #accession A94284

#molecule-type mRNA
 #residues 1-595 #label GR1
 #cross-references GB:M12674; NID:g182192; PID:g182193
 REFERENCE A93376
 #authors Argos, S.; Walter, P.; Kumar, V.; Krust, A.; Bornert, J.M.;
 #journal Nature (1986) 320:134-139
 #title Human estrogen receptor cDNA: sequence, expression and
 homology to v-erb-A.
 #cross-references MUID:86146892
 #accession A93376
 #molecule-type mRNA
 #residues 1-595 #label GR2
 #cross-references GB:X03635; NID:g31233; PID:g31234
 REFERENCE A43021
 #authors Ponglikitmongkol, M.; Green, S.; Chambon, P.
 #journal EMBO J. (1988) 7:3385-3388
 #title Genomic organization of the human oestrogen receptor gene.
 #accession A43021
 #molecule-type DNA
 #residues 143-161;205-225;244-264;356-374,'G',376;402-422;447-460,
 'P',462-467;508-528 #label PON
 #cross-references EMBL:X62462; NID:g31201; PID:g31205
 REFERENCE S27140
 #authors Keaveney, M.; Klug, J.; Gannon, F.
 #journal DNA Seq. (1992) 2:347-358
 #title Sequence analysis of the 5' flanking region of the human
 estrogen receptor gene.
 #accession S27143
 #status preliminary; translation not shown
 #molecule-type DNA
 #residues 1-115 #label KEA
 #cross-references EMBL:X62462; NID:g31201; PID:g31205
 REFERENCE S34000
 #authors Pfeiffer, U.; Pecarotta, F.; Castagnetta, L.; Vidali, G.
 #journal Cancer Res. (1993) 53:741-743
 #title Estrogen receptor variant messenger RNA lacking exon 4 in
 estrogen-responsive human breast cancer cell lines.
 #accession S34000
 #status preliminary
 #molecule-type mRNA
 #residues 216-254,367-399,'G',401-434 #label PFE
 #cross-references EMBL:X73067; NID:g579865; PID:g939886
 #note the authors translated the codon GGC for residue 400 as
 Val
 REFERENCE A41925
 #authors Dotzlaw, H.; Alkhalaf, M.; Murphy, L.C.
 #journal Mol. Endocrinol. (1992) 6:773-785
 #title Characterization of estrogen receptor variant mRNAs from
 human breast cancers.
 #cross-references MUID:92293154
 #accession A41925
 #molecule-type mRNA
 #residues 1-214,'ELPTLC' #label DOT
 #cross-references GB:M69297; NID:g182218; PID:g182219
 #note sequence has been revised after extraction from NCBI
 #experimental_source clone 4; Breast cancer
 #note sequence has been revised after extraction from NCBI
 #note backbone
 #note the complete sequence of neither the nucleotide nor the
 protein is shown in this paper
 #note sequence extracted from NCBI backbone (NCBI:106580)
 #accession B41925
 #status significant sequence differences
 #molecule-type mRNA
 #cross-references GB:M69296
 #experimental_source clone 24; Breast cancer
 #note sequence extracted from NCBI backbone (NCBI:106597)
 #note regulation of eukaryotic gene expression and affect cellular
 proliferation and differentiation in target tissues.
 #note In the absence of 19and, steroid hormone receptors are thought to
 be weakly associated with nuclear components; hormone binding
 greatly increases receptor affinity. The hormone-receptor complex
 appears to recognize discrete DNA sequences upstream of

[illegible]

QY	474	EHLNMGCKNVVPPYDLELNLNHNHVLARGCKSSITGSECPADSKSRESQNPQS	529
RESULT	7		
ENTRY	147140	#type complete	
TITLE	estradiol receptor - pig		
ORGANISM	#formal_name Sus scrofa domestica	#common_name domestic pig	
DATE	21-Feb-1997	#sequence_revision 21-Feb-1997	#text_change
ACCESSIONS	147140		
REFERENCE	147140		
#authors	Bokenkamp, D.; Jungblut, P. W.; Thole, H. H.		
#journal	Mol. Cell. Endocrinol. (1994) 104:163-172		
#title	The C-terminal half of the porcine estradiol receptor contains no post-translational modification: determination of the primary structure.		
#cross-references	MU01:95080454		
#accession	I47140		
##status	preliminary: translated from GH/EMBL/DBD		
##molecule_type	mRNA		
##residues	1-595	#label BOK	
##cross-references	EMBL:Z37167; NID:G587554; PID:G587555		
REFERENCE	S66250		
#authors	Thole, H. H.; Maschler, I.; Jungblut, P. W.		
#journal	Eur. J. Biochem. (1995) 231:510-516		
#title	Surface mapping of the ligand-filled C-terminal half of the porcine estradiol receptor by restricted proteolysis.		
#accession	S66250		
##status	preliminary		
##molecule_type	protein		
##residues	237-307; 310-313; 320-323; 329-332; 337-340; 417-420; 466-473		
##label	THO		
CLASSIFICATION	#superfamily estrogen receptor; erba transforming protein homology		
KEYWORDS	steroid hormone receptor; zinc finger		
FEATURE			
183-456	#domain erba transforming protein homology		
SUMMARY	#length 595 #molecular_weight 66361 #checksum 5521		
Query Match	42.4% Score 1654; DB 2; Length 595;		
Best Local Similarity	53.9% Pred. No. 5.58e-303;		
Matches	242; Conservative 97; Mismatches 90; Indels 20; Gaps 15;		
Db	104 SPSPVLVLPHPOLSPPLHPGQOVPPYLLENESGY-AVREACPAPAFYRPSNDRRQGR	162	
QY	73 TTSNVNLPMPGTHLSP-LVYHROLSHYLAPQKSPMCWEARSLSHLTPVYNETLAKRYSG-	130	
Db	163 ERLASTSDKGSAMASAKETRYCAVCNDVASYGHYGVASCCEGKAPFKRSIOGHNDYMP	222	
QY	131 NRCASPV-TGP-G--SKRDAHFCVSCDYASGYHGVWSCGCKAPFKRSIOGHNDYICP	186	
Db	223 ATNCTIDKNRKCKQACRLKCEVGMAMGGIRKDRGRGRLMKHKRQDRDGGREANVP	282	
QY	187 ATNCTIDKNRKCKQACRLKCEVGMAMGGIRKDRGRGRLMKHKRQDRDGGREANVP	282	
Db	283 PGDMRSANLWSPPLILHTKKNPSVLSLTFADOMISALTEAPPLIYSEVPTPLSEASM	342	
QY	244 KA-KRSGG--HAPR-VNELLDA--LS-PEQVLVTLLEPAPHVLIS-RPSAPFTFASH	294	
Db	343 MGLTLNLADRELVMINMARVPGFDLSLHDVNLLECAMLEIMIGLWRSMEHPGKL	402	
QY	295 MMSLTKLADRELVMINMARVPGFDLSLHDVNLLECAMLEIMIGLWRSMEHPGKL	354	
Db	403 LFAANLLDRQKCGVAGWEIPDMLATSSRRPMALJOSEEVVCLKSTLLNSGYTFL	462	
QY	355 IFADVLVLDREGCGVAGWEIPDMLATSSRRPMALJOSEEVVCLKSTLLNSGYTFL	414	
Db	463 SSTLSLEEKHHRVAVLTDITLHLMATAAGLLOQHRLAKQLLISHFPMNSKGM	522	
QY	415 TATODA-DSSKRLAHLLNATVDALVAVYIAKSGISSQOQSHRLANMLLSHVHASNKGK	473	
Db	523 EHLNMGCKNVVPPYDLELNLNHNHVLARGCKSSITGSECPADSKSRESQNPQS	529	

Matches	140;	Conservative	84;	Mismatches	108;	Indels	22;	Gaps	14
Db	101	RLCLVGDIASGHHYGYVASC	EACCAKAFKRTIIOGNI	ESCPATNECETITRRKRSQ	ACRF	160			
QY	147	HFCVACSDYASGHHYGYVASC	EGCCKAFKFRSIOGHNDYIC	PATNOCTIDKNRKRKSCQ	ACRL	206			
Db	161	MKCLAVGMLKEGVADLVY--R	-GGKQRY-KRRLDSENSPL	LSQISPPAKPPLTIYS--	214				
QY	207	RKCEYGVAMVKGSGSRREK	CGRYLVNRQSSADEQLHC	ACAKARSGGHARVVELLDAL	SP	266			
Db	215	YL-L-VAEPDKLYAM-PPD	VPEDIDKATLTTCIDLA	REVLFIEMNAKHIPGSE	NLTIGD	271			
QY	267	QLVLTLEAEPHYLIRSP	SAPFLFASMMSTYKTLAD	KELVHMISMAKKIPGFVEL	SFLD	326			
Db	272	QMSLLOSAMEILLIGIVY	SLPYDDKLAAYEDYIMDE	HSRLV-GLLELYRAILQ	LVRR	330			
QY	327	QVRLESCMWEVLMMGM	WMRSIDHPGKLIAPDVL	DRDQKCEGILEIFDM	LATTSR	386			
Db	331	YKTKLKEKEEYVMIKALA	LANDSMY--T---ENLE	VOKLQDLH--BALQ	DYLS-	380			
QY	387	FRELKLOHKEILCYKAM	ILNTS-SMYPVLVATOP	ADSSRKLAHLINAVTAL	VWVIKSG	445			
Db	381	-QRHEEDRRAGKLLTLP	LRQTAARVONHYSVK	LOGKVPMMHLEFEM	LEAKV	433			
QY	446	ISSQOQSRRLANLML	LSHVRRHASKMEHLN	CKNVPYIDLTLEM	NAHV	499			
RESULT	12								
ENTRY		S58087	#type complete						
TITLE			estrogen receptor related protein - mouse						
ORGANISM			#format_name Mus musculus	#common_name house mouse					
DATE			13-Jan-1996	#sequence_revision 01-Mar-1996	#text_change 31-Oct-1997				
ACCESSIONS		S58087							
REFERENCE		S58087							
#authors			Petersson, K.; Svensson, K.; Mattsson, R.; Carlsson, B.; Ohlsson, R.; Berkenstam, A.						
#submission			submitted to the EMBL Data Library, July 1995						
#description			Expression of a novel member of estrogen response element-binding nuclear receptors is restricted to the early stages of chorion formation during mouse embryogenesis.						
#accession		S58087							
#status			preliminary						
##molecule_type		mRNA							
##residues		1-433	#label PET						
##cross-references		EMBL:X89594							
CLASSIFICATION			#superfamily unassigned	erba-related proteins; erba transforming protein homology steroid hormone receptor; zinc finger					
KEYWORDS									
FEATURE									
101-352									
SUMMARY			#domain	erba transforming protein homology	#label ERBA				
			#length 433	#molecular-weight 48400	#checksum 1075				
Query Match			20.5%;	Score 799;	DB 2;	Length 433;			
Best Local Similarity			38.7%;	Pred. No. 1.97e-128;					
Matches	137;	Conservative	85;	Mismatches 110;	Indels 22;	Gaps 15;			
Db	101	RLCLVGDIASGHHYGYVASC	EACCAKAFKRTIIOGNI	ESCPATNECETITRRKRSQ	ACRF	160			
QY	147	HFCVACSDYASGHHYGYVASC	EGCCKAFKFRSIOGHNDYIC	PATNOCTIDKNRKRKSCQ	ACRL	206			
Db	161	MKCLAVGMLKEGVADLVY--R	-GGKQRY-KRRLDSENSPL	LSQISPPAKPPLTIYS-N	215				
QY	207	RKCEYGVAMVKGSGSRREK	CGRYLVNRQSSADEQLHC	ACAKARSGGHARVVELLDAL	SP	266			
Db	215	YL-L-VAEPDKLYAM-PPND	IPEDIDKATLTTCIDLA	REVLFIEMNAKHIPGSE	NLTIGD	271			
QY	267	QLVLTLEAEPHYLIRSP	SAPFLFASMMSTYKTLAD	KELVHMISMAKKIPGFVEL	SFLD	326			
Db	272	QMSLLOSAMEILLIGIVY	SLPYDDKLAAYEDYIMDE	HSRLV-GLDDYRAILQ	LVRR	330			
QY	327	QVRLESCMWEVLMMGM	WMRSIDHPGKLIAPDVL	DRDQKCEGILEIFDM	LATTSR	386			
Db	331	YKTKLKEKEEYVMIKALA	LANDSMY--T---ENLE	VOKLQDLH--BALQ	DYLS-	380			
QY	387	FRELKLOHKEILCYKAM	ILNTS-SMYPVLVATOP	ADSSRKLAHLINAVTAL	VWVIKSG	445			
Db	381	-QRHEEDRRAGKLLTLP	LRQTAARVONHYSVK	LOGKVPMMHLEFEM	LEAKV	433			
QY	446	ISSQOQSRRLANLML	LSHVRRHASKMEHLN	CKNVPYIDLTLEM	NAHV	499			

Db	331	YKLLVKEKEFFILKRLATLANSYSM--I---ENLEAVQLODQLH---EALQ-DVETLS	380
Oy	367	FRELKQKHREYLCVARMILLNS-SMPLPYTAOTADADSSKRLAHLNAYDALVWVIAKSG	445
Db	361	-QRHEEPFRAGKILWTLPLRLQRTAAKAVOHFYSVKLOGKVMKRLPLEMLEARV	433
Oy	446	ISSQOOSMRLANILMLLSHVRRASNKGMEHLNMKCKNVYPVYDLLEMLNAVH	499
RESULT	13		
ENTRY			
TITLE	A29345	#type complete	
ALTERNATE_NAMES	steroid hormone receptor ERRI precursor - human		
ORGANISM	estrogen-related receptor		
DATE	#formal_name Homo sapiens	#common_name man	
	15-Dec-1988	#sequence_revision 15-Dec-1988	#text_change
	12-Sep-1997		
ACCESSIONS	A29345; A49074		
REFERENCE	A29345		
#authors	Giguere, V.; Yang, N.; Segui, P.; Evans, R.M.		
#journal	Nature (1988) 331:91-94		
#title	Identification of a new class of steroid hormone receptors.		
#cross-references	MUID:88122546		
#accession	A29345		
#molecule_type	mRNA		
#residues	1-521	#label GIG	
#cross-references	EMBL:X51416; NID:g36608; PID:g36609; EMBL:Y00290		
REFERENCE	A49074		
#authors	Wiley, S.R.; Kraus, R.J.; Zuo, F.; Murray, E.E.; Loritz, K.; Mertz, J.E.		
#journal	Genes Dev. (1993) 7:2206-2219		
#title	Sv40 early-to-late switch involves titration of cellular transcriptional repressors.		
#accession	A49074		
#status	preliminary		
#molecule_type	protein		
#residues	166-169, 'X', 171-173	#label WIL	
CLASSIFICATION	#superfamily unassigned	estrogen-related proteins; estrogen-related protein	
KEYWORDS	DNA binding; nucleus; steroid hormone receptor; transcription regulation; zinc finger		
FEATURE			
174-434	#domain	estrogen transforming protein	homology #label ERBA\
176-196	#region	zinc finger	
212-236	#region	zinc finger	
SUMMARY	#length 521	#molecular_weight 55742	#checksum 5840
Query Match	20.5%	Score 800.	DB 2; Length 521;
Best Local Similarity	39.98%	Pred. No. 1.24e-128;	
Matches 141;	Conservative	74;	Mismatches 128; Indels 10; Gaps 7;
Db	174	RLLCTCGDVASVYHGVACCEACAKFAFKRTIOGSEICSPASNECETTKRRKACQCRF	233
Oy	147	HFCVASCSDPASGTHYVSWGCEGKAFKFSKISGHNDYICATQCTIDTKRRKSCQACRL	206
Db	234	TKCLRNGLKEGVRLDVRBGROKYKRRPEVDLPFRFG-P-PRAGPLAAGGPKTAAPV	291
Oy	207	RKCYEGVMWKCGRSRRRC-GYRLVRQRQSADEQLHCAGKAKKRGSHAPRVRELLDALSP	265
Db	292	NALVSHLVVEEBEKLKAMDPDAGPDSCHLRAVALTLCDFPRELYVTVTISMAKSTGFSSLSL	351
Oy	266	EDLVLTLLAEPRPHV-L-ISRPSAPFTEDASMMSLTKLADKELVHMSWARKIGFVELSL	324
Db	352	SDQMSVGVNMEVILVAGVORSLPQDELDAFADLYLD-EEGARAAGLGLGALLQLV	410
Oy	325	PQVALLBECWMEVILMGLMMSIDPRGLKIFRPDLVLDDEKCYEGVILFIDMLATT	384
Db	411	RRLQALRLREBEYVILKALALANSDSVNIHDEDRPLWSCEKLLN--EAL--LEYGRA	465
Oy	385	SRFRELKQKHREYLCVARMILLNS-SMPLPYTAOTADADSSKRLAHLNAYDALVWVIAKS	444
Db	466	GGGGAERRRAGRLTLTLPRLQRTAAKAVLANHFGVNLKGVKVPYRNKFLFLEMLA	518

OY	445	GISSQOOSMRLANLMLSHVRASAKGGEHLNKKCKNVVPVYLLEMLNA	497
RESULT	14	A39596 #type complete	
ENTRY		progesterone receptor B form - mouse	
TITLE		#formal_name Mus musculus #common_name house mouse	
ORGANISM		20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change	
DATE		31-Oct-1997	
ACCESSIONS		A39596	
REFERENCE		A39596	
#authors		Schott, D.R.; Shyamala, G.; Schneider, W.; Parry, G.	
#journal		Biochemistry (1991) 30:7014-7020	
#title		Molecular cloning, sequence analyses, and expression of complementary DNA encoding murine progesterone receptor.	
#cross-references		complementary DNA encoding murine progesterone receptor.	
#accession		K01D:9129/759	
#status		A39596	
##molecule_type		mRNA	
##residues		1-923 ##label SCH	
##cross-references		GB:M68915; GB:J05333; NID:g200471; PID:g200472	
CLASSIFICATION		#superfamily progesterone receptor; erba transforming protein homology	
KEYWORDS		DNA binding; nucleus; steroid hormone receptor; transcription regulation; zinc finger	
FEATURE			
555-819		#domain erba transforming protein homology #label ERBA\	
557-577		#region zinc finger	
593-617		#region zinc finger	
SUMMARY		#length 923 #molecular-weight 99073 #checksum 6627	
Query Match		17.4%; Score 681; DB 2; Length 923;	
Best Local Similarity		30.8%; Prid. No. 6,28e-105;	
Matches		123; Conservative 112; Mismatches 138; Indels 27; Gaps 27;	
Db	491	AAPGTAAPIATYOPLG-LINGLPOLG-Y-QAAVLKDSLPQVPEPYLYLRPDSEASOSPQY	547
OY	81	PTPGHLSPLVYHHQLSHLVAEPQSKWCARSLERTLP-VNRETLK-RKVSNGRCASFVT	138
Db	548	GFDSDPQ-KICLICGDSEASCHGYLTGCSCVFRRAMEGOHNLYLCAGRDCIYDKIR	606
OY	139	GGSGKRDAHFCAVCSYPASGYHGVMSCEGCAPFRKSIOGHNDVICPATNOCTIDKNRR	198
Db	607	KNCFACILKRCQCAQAWL-GGRKF-KFNKYVMKTLGVALPOSGLPNESQALSORTIT	664
OY	199	KSCACALRCFYCVGWKCGSRERRCGYLTRORSADP-QL-HCAGRAKRSQGHAPVR	256
Db	665	FSPNQELQWP-ELINLMNSIEPDVYYAHDNTKRPDTSSILTSLNQGEROLISYKWS	723
OY	257	-ELLIDA-LSPEDLVLTLEAEPPHYLLSRPSA-PFTESMMMSLTKLADKLVAHMISWA	313
Db	724	KSLPGFNLIIIDOITLIQYSWMSLVFGLGSRYSKHVSGOMLYEAPDLILN-EQ-R-MK	780
OY	314	KRIGFVELSLDFQVRLBSCKWEVYLMGMIMRSIDH-PGKIL-FAPDLVDREBGCYE	371
Db	781	E-LSFISCLTMMQIQOEYFKIQVTHEBFLCKAVILLANT-I-PLDGIRSOGFEEMRSS	837
OY	372	GILEIFDM-L-L-ATTSRFRELTKHQHEKLCYKAMILLNSSMYPLVATQADSSRKLAH	428
Db	838	YIRELIKAIQ-LROKGVPISORFYTLTLLDSLMDYKQ	876
OY	429	LNNVTDALVWIAKGISQQOOSMRLANL-MLLSHVH	467
RESULT	15	I53280 #type complete	
ENTRY		progesterone receptor - rat	
TITLE		#formal_name Rattus norvegicus #common_name Norway rat	
ORGANISM		02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change	
DATE		25-Apr-1997	
ACCESSIONS		I53280; B23733; A23733	
REFERENCE		I53280	
#authors		Park-Sarge, O.K.; Mayo, K.E.	

THIS PAGE BLANK (USPTO)


```

QY 61 NTYNLEGGGROTTSPNVLMPTPGHLSPLVYHROLSHYAEPOKSPWCARSLEHTLPVN 120
DB 140 RETLKRKLSGSSCASPVTPSPNARKDAHFCVACSDYASGYHYGWPCEGCKAFKRSIOGH 199
QY 121 RETLKRKVSNGNRCAAPVTGSGKRDHAFCAVCSDYASGYHYGWPCEGCKAFKRSIOGH 180
DB 200 NDYICPATNOCTIDKRRKRSQACRLRKCEYGVGMVCGSRRERCGYRLVRRORSSEQYH 259
QY 181 NYICPATNOCTIDKRRKRSQACRLRKCEYGVGMVCGSRRERCGYRLVRRORSSEQYH 240
DB 260 CLSKARNGSHAPRVKELLSTLSPQVLYTLLEAPPNVLSRRPMPTFEASMMMSLTK 319
QY 241 CAGKAKRSNGHAPRVKELLSTLSPQVLYTLLEAPPNVLSRRPMPTFEASMMMSLTK 300
DB 320 LADKELVHMISWAKKIPGFVELSLDQVRLLESCMVEYLMGMMSIDHPGKLIFAPDL 379
QY 301 LADKELVHMISWAKKIPGFVELSLDQVRLLESCMVEYLMGMMSIDHPGKLIFAPDL 360
DB 380 VLDROGKCVESGILEIFDMMLATTSRFRKLQHKREYLCVAMILLNSSMYPLVATODA 439
QY 361 VLDROGKCVESGILEIFDMMLATTSRFRKLQHKREYLCVAMILLNSSMYPLVATODA 420
DB 440 ESSRKTLHLNAYTDALVWYIAKSGISSQOOSMRLANLMLLSHVHSHKNGEHLISMK 499
QY 421 DSRKLAHLNAYTDALVWYIAKSGISSQOOSMRLANLMLLSHVHSHKNGEHLISMK 480
DB 500 CKNVVYVDLLEMLNAHLRGYKPSISGSECSSTEDSKNKSQONLQSQ 549
QY 481 CKNVVYVDLLEMLNAHLRGYKPSISGSECSSTEDSKNKSQONLQSQ 530

```

RESULT 2
ID Q92731 PRELIMINARY: PRT: 477 AA.

```

AC 092731;
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE ESTROGEN RECEPTOR BETA.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-TESTIS.
RA MOSELMAN S., POLMAN J., DIKREMA R.;
RA MEDLINE: 96354875.
RL FEBS LETT. 392:49-53(1996).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: X99101; E255390; -.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR. 1.
DR PFM: PF00104; hormone_rec.
DR PFM: PF00105; zf-C4.
KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW ZINC-FINGER.
SQ SEQUENCE 477 AA: 5383 MW; 4E76E987 CRC32;

```

Query Match 90.5%; Score 3532; DB 4; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

DB 1 MNYISPSNVTNLEGGGROTTSPNVLMPTPGHLSPLVYHROLSHYAEPOKSPWCARSLE 60
QY 54 MNYISPSNVTNLEGGGROTTSPNVLMPTPGHLSPLVYHROLSHYAEPOKSPWCARSLE 113
DB 61 EHTLPNRETLKRVSGNRCASPVTPGSKRDHAFCAVCSDYASGYHYGWPCEGCKAF 120
QY 114 EHTLPNRETLKRVSGNRCASPVTPGSKRDHAFCAVCSDYASGYHYGWPCEGCKAF 173
DB 121 KRSIOGHNDYICPATNOCTIDKRRKRSQACRLRKCEYGVGMVCGSRRERCGYRLVRROR 180
QY 174 KRSIOGHNDYICPATNOCTIDKRRKRSQACRLRKCEYGVGMVCGSRRERCGYRLVRROR 233
DB 181 SADBOLHCAGKAKRSNGHAPRVKELLSTLSPQVLYTLLEAPPNVLSRRPMPTFEAS 240

```

```

QY 234 SADBOLHCAGKAKRSNGHAPRVKELLSTLSPQVLYTLLEAPPNVLSRRPMPTFEAS 293
DB 241 MMSLTKLADKELVHMISWAKKIPGFVELSLDQVRLLESCMVEYLMGMMSIDHPGK 300
QY 294 MMSLTKLADKELVHMISWAKKIPGFVELSLDQVRLLESCMVEYLMGMMSIDHPGK 353
DB 301 LIFAPDLVLDROGKCVESGILEIFDMMLATTSRFRKLQHKREYLCVAMILLNSSMYPL 360
QY 354 LIFAPDLVLDROGKCVESGILEIFDMMLATTSRFRKLQHKREYLCVAMILLNSSMYPL 413
DB 361 VTATODADSSRLAHLNAYTDALVWYIAKSGISSQOOSMRLANLMLLSHVHSHKNGM 420
QY 414 VTATODADSSRLAHLNAYTDALVWYIAKSGISSQOOSMRLANLMLLSHVHSHKNGM 473
DB 421 EHLNKKCKNVVYVDLLEMLNAHLRGYKPSISGSECSSTEDSKNKSQONLQSQ 477
QY 474 EHLNKKCKNVVYVDLLEMLNAHLRGYKPSISGSECSSTEDSKNKSQONLQSQ 530

```

RESULT 3
ID O60685 PRELIMINARY: PRT: 513 AA.

```

AC O60685;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE ESTROGEN RECEPTOR BETA 3 ISOFORM.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA MOORE J.T., MCKEE D.D., SLENTZ-KESLER K., MOORE L.B., JONES S.A.,
RA SU J.L., HORNE E.L., KLEWER S.A., LEHMAN J.M., WILLSON T.M.;
RL SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: AF060555; G3091286; -.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR. 1.
KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW ZINC-FINGER.
SQ SEQUENCE 513 AA: 57518 MW; 1CC02BEC CRC32;

```

Query Match 89.1%; Score 3480; DB 4; Length 513;
Best Local Similarity 97.7%; Pred. No. 0.00e+00;
Matches 470; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

```

DB 1 MDIKNPSLSPSSSYNCSOSILPLEHGSITYPPSYVDSHHEYPAMTYSPAVMNYSTPS 60
QY 1 MDIKNPSLSPSSSYNCSOSILPLEHGSITYPPSYVDSHHEYPAMTYSPAVMNYSTPS 60
DB 61 NTYNLEGGGROTTSPNVLMPTPGHLSPLVYHROLSHYAEPOKSPWCARSLEHTLPVN 120
QY 61 NTYNLEGGGROTTSPNVLMPTPGHLSPLVYHROLSHYAEPOKSPWCARSLEHTLPVN 120
DB 121 RETLKRKVSNGNRCAAPVTGSGKRDHAFCAVCSDYASGYHYGWPCEGCKAFKRSIOGH 180
QY 121 RETLKRKVSNGNRCAAPVTGSGKRDHAFCAVCSDYASGYHYGWPCEGCKAFKRSIOGH 180
DB 181 NDYICPATNOCTIDKRRKRSQACRLRKCEYGVGMVCGSRRERCGYRLVRRORSSEQYH 240
QY 181 NDYICPATNOCTIDKRRKRSQACRLRKCEYGVGMVCGSRRERCGYRLVRRORSSEQYH 240
DB 241 CAGKAKRSNGHAPRVKELLSTLSPQVLYTLLEAPPNVLSRRPMPTFEASMMMSLTK 300
QY 241 CAGKAKRSNGHAPRVKELLSTLSPQVLYTLLEAPPNVLSRRPMPTFEASMMMSLTK 300
DB 301 LADKELVHMISWAKKIPGFVELSLDQVRLLESCMVEYLMGMMSIDHPGKLIFAPDL 360
QY 301 LADKELVHMISWAKKIPGFVELSLDQVRLLESCMVEYLMGMMSIDHPGKLIFAPDL 360
DB 361 VLDROGKCVESGILEIFDMMLATTSRFRKLQHKREYLCVAMILLNSSMYPLVATODA 420
QY 361 VLDROGKCVESGILEIFDMMLATTSRFRKLQHKREYLCVAMILLNSSMYPLVATODA 420

```



```

DN      01-JUL-1997 (TREMBLREL. 04, CREATED)
DT      01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DI      01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE      ESTROGEN RECEPTOR BETA.
EN      ESTRB.
GN      MUS MUSCULUS (MOUSE).
OS      MUSCULUS (MOUSE).
OC      EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC      EUETHERIA; RODENTIA.
RN      [1]
RP      SEQUENCE FROM N.A..
RC      STRAIN=129/SV; TISSUE=OVARY;
RA      TREMBLAY G.B., TREMBLAY A., COPELAND N.G., GILBERT D.J., JENKINS N.A.,
RA      LABRIE F., GAUDIERE V.;
RL      NATURE 11:353-365(1977).
CC      -1 SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR      EMBL; U01451; G1912468; -.
DR      MGD; MG1:109392; ESTRB.
DR      PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
DR      PFAM; PF00104; hormone_rec.
DR      RFAM; RF00105; zf-C4.
KW      RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW      ZINC-FINGER.
SQ      SEQUENCE 485 AA; 54058 MW; B6FBE7C7 CRC32;
DB      Query Match          84.0%; Score 3278; DB 11; Length 485;
      Best Local Similarity 88.9%; Pred. No. 0.00e+00;
      Matches 431; Conservative 34; Mismatches 20; Indels 0; Gaps 0
Dy      1 MTFSPAVMNTSVSSSGNLEGGVYROTASFNVLMPISGHSIPLATQCSSLTIAEPDKS 60
      |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
      46 MTFESPVMNNTSIIPSNVTNEGEGGROTTSPNVLMPTPGHLSPLVYHROLSHLTAAEPDKS 105

```

[illegible]

CC	EUHAROTIA: METAFOA; CHORDATA: VERTEBRATA; TETRAPODA; MAMMALIA;
CC	EUHAROTIA: ROENTHIA.
CC	[1]
CC	SEQUENCE FROM N.A.
CC	STRAIN-SPRAGUE DAWLEY; TISSUE=PROSTATE;
CC	RA KUIPER G.G.J.M., ENMARK E., PELTO-HOIKKO M., NILSSON S.,
CC	RA GUSTAFSSON J.A.;
CC	PROT. NATL. ACAD. SCI. U.S.A. 93:5925-5930(1996).
CC	-1C SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC	EMBL: U57439; G1373281; .
CC	DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
CC	DR PFAM; PF00104; hormone_rec.
CC	DR PFAM; PF00105; zf-C4.
CC	DR RECEPTOR: TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
CC	ZINC-FINGER.
CC	SEQUENCE 485 AA; 54199 MW; BAF55792 CRC32;
CC	Query Match 83.9%; Score 3274; DB 11; Length 485;
CC	Best Local Similarity 88.7%; Pred. No. 0.00e+00;
CC	Matches 430; Conservative 35; Mismatches 20; Indels 0; Gaps 0
CC	Db 1 MTFSPAVMANTSVPGTSLNDGCPVRLSTSPNVLMPTISGLHSPLATHCQSSLYAEPQKS 60
CC	46 MTFEYPAVMNYSIPSNVNLTEGGPQRTSPNVLMPTPGHLSPLVHROLSHYAEPQKS 105
CC	Db 61 PWCERSLIEHTLPVARETLKRLKSSSCASASTPSMARNDHFCPCSDYASGHHYQWS 120
CC	106 PWCERSLIEHTLPVARETLKRLKSSGNCASPVTPGGSKRDAHFCAYCSDYASGHHYQWS 165
CC	Db 121 CEGCAAFKRSIOGNDYICPATNOCTIDKNRRKSCQACRLKCYEYGVNVCGRRCRG 180
CC	166 CEGCAAFKRSIOGNDYICPATNOCTIDKNRRKSCQACRLKCYEYGVNVCGRRCRG 225
CC	Db 181 YRIYRORSSEQVHCLSKARNRNGHAPVARELLSTSPQVLTLLAEPPNVLSRP 240
CC	226 YRIYRORSADQHLCAKAKRSGHAPVARELLDALSPQVLTLLAEPPNVLSRP 285
CC	Db 241 SMPFEASMMSTLTKADELVNMIQNAKKIIPGEYELSLDQVRLSCMEVYLVGLMW 300
CC	286 SAPFEASMMSTLTKADELVNMIQNAKKIIPGEYELSLDQVRLSCMEVYLVGLMW 345
CC	Db 301 RSIDHPGKLTAPDVLDRDQKCYEGILETFDMTLATTSRFRKLQHKREYLCVKAMIL 360
CC	346 RSIDHPGKLTAPDVLDRDQKCYEGILETFDMTLATTSRFRKLQHKREYLCVKAMIL 405
CC	Db 361 LNSSMYPLASANOEAESSRKLTHLNAATYDALVWIAKSGISSQOQSVRLANILMLTSHV 420
CC	406 LNSSMYPLVTAQDADSSRKLHLNANVTDALVWIAKSGISSQOQSVRLANILMLTSHV 465
CC	Db 421 RHISKGEHLISKCKNVVPPYDILLMLNANHILRGYKSSISGSEGSSTEDSKNKSSQ 480
CC	466 RHASKGEHLILMKCKNVVPPYDILLMLNANHILRGYKSSISGSEGSSTEDSKNKSSQ 525
CC	Db 481 NLSQ 485
CC	526 NPSQ 530
CC	RESULT 7
CC	ID 035635 PRELIMINARY: PRT: 485 AA.
CC	AC 035635.
CC	DT 01-JAN-1998 (TREMBLREL.05, CREATED)
CC	DT 01-JAN-1998 (TREMBLREL.05, LAST SEQUENCE UPDATE)
CC	DT 01-AUG-1998 (TREMBLREL.07, LAST ANNOTATION UPDATE)
CC	DE ESTROGEN RECEPTOR, SUBTYPE BETA.
CC	OS MUS MUSCULUS (MOUSE).
CC	OC EUHAROTIA: METAFOA; CHORDATA: VERTEBRATA; TETRAPODA; MAMMALIA;
CC	OC EUHAROTIA: ROENTHIA.
CC	CC [1]
CC	CC SEQUENCE FROM N.A.
CC	CC PETERSSON K., GRANDIEN K., KUIPER G.G., GUSTAFSSON J.A.;
CC	CC SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC	CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

Query Match	83.2%	Score 3250;	DB 11;	Length 485;
Best Local Similarity	88.0%;	Pred. No. 0.00e+00;		
Matches 427;	Conservative	36;	Mismatches 22;	Indels 0;
			Gaps	0;
Db	1	MAFSPAYMANYSVSSSTGNLEGGPVROTASPNVLMPIRSGHLSPLATHCOSSLLYAEPOKS	60	
Qy	46	MTFTSPAYMANYSISVNTNLEGGRGRTTSPNVLMPFPHLSPLVYHROLSHLAEPOKS	105	
Db	61	PWCARSLEHLPVNRRETLRKLKLGSGSCASPVTSPTKRPAHFCAMVCSDAASGHHYGWS	120	
Qy	106	PWCARSLEHLPVNRRETLRKLKLGSGSCASPVTSPTKRPAHFCAMVCSDAASGHHYGWS	165	
Db	121	CEGKAPFKRSIQGHNDYICPATNQCTTIDKNRKNCOACRLKCYEYGVYKCSRRERCG	180	
Qy	166	CEGKAPFKRSIQGHNDYICPATNQCTTIDKNRKNCOACRLKCYEYGVYKCSRRERCG	225	
Db	181	YRIRPQRSASEOYCHLKNKAKRTSGHPRYKELLNSLSPQVLVTLLEAPPVNLVSRP	240	
Qy	226	YRIRPQRSASEOYCHLKNKAKRTSGHPRYKELLNSLSPQVLVTLLEAPPVNLVSRP	285	
Db	241	SMPTFASMMMSLTKLADKELVHIGAKKIPGEVELSLDDOYVLLSECMVEYLVAGLAW	300	
Qy	286	SAPFTFASMMMSLTKLADKELVHIGAKKIPGEVELSLDDOYVLLSECMVEYLVAGLAW	345	
Db	301	RSIDHPGKLIFAPDLVLDREBGKCVESILETFDMLLATARFRELKLOHKEVLCVKAMIL	360	
Qy	346	RSIDHPGKLIFAPDLVLDREBGKCVESILETFDMLLATARFRELKLOHKEVLCVKAMIL	405	
Db	361	LNSSYTHLATSQSAESSRKLTHLNLAVTDALVWIVKSKRISQOQSVRLANLMLLSHV	420	
Qy	406	LNSSYTHLATSQSAESSRKLTHLNLAVTDALVWIVKSKRISQOQSVRLANLMLLSHV	465	
Db	421	RHISNKGHEHLLSMCKKNVYVYDLLLEMLNAHTLRGKSSISGSGCSTEDSKSKSGSQ	480	
Qy	466	RHISNKGHEHLLSMCKKNVYVYDLLLEMLNAHTLRGKSSISGSGCSTEDSKSKSGSQ	525	
Db	481	NLQSG 485		
Qy	526	NPOSG 530		
RESULT	8	PRELIMINARY;	PRT;	503 AA.
ID	035785			
AC	035785			
DT	01-JAN-1998	(TREMBL.REL. 05, CREATED)		
DT	01-JAN-1998	(TREMBL.REL. 05, LAST SEQUENCE UPDATE)		
DT	01-JUN-1998	(TREMBL.REL. 06, LAST ANNOTATION UPDATE)		
DE	OESTROGEN RECEPTOR BETA2.			
GN	ERBETA.			
OS	RATTUS NORVEGICUS (RAT).			
CC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
CC	EUTHERIA; RODENTIA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	STRAIN-WISTAR; TISSUE--PROSTATE;			
RA	ALDRIDGE T.C.;			
RL	SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RP	STRAIN-SPRAGUE-DAWLEY;			
RA	PETERSEN D.N., TKALCEVIC G.T., KOZA-TAYLOR P.H., TURI T.G.,			
RL	BROWN T.A.;			
RL	ENDOCRINOLOGY 0:0-0(1998).			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).			
CC	EMBL; AJ002603; E1184907; -.			


```

QY 106 PWCBAFSLHTLPVNRKTLKRVSGNRCASVYTGPGSKRDHFCACVSDYASGYHGWMS 165
DB 121 CEGCKAFKRSIQG-----SRRERC 141
QY 166 CEGCKAFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLKCYEYGMVCKGSRRCRG 225
DB 142 YRIVRORSSSEVYCHCSKAKRNGGAPRVKELLSTLSPQVLVTLLEAPPVLSRP 201
QY 226 YRIVRORSADEQLHCAGKAKRSRGHAPRVRELLDLALSPQVLVTLLEAPPVLSRP 285
DB 202 SMPTEASMMASLTKLADKELVHMIGMAKKIPGFVELSLDQVRLSCMVEVLMGLMW 261
QY 286 SAPTEASMMASLTKLADKELVHMISMAKKIPGFVELSLDQVRLSCMVEVLMGLMW 345
DB 262 RSIDHPGKLIAPDVLVDRDEGKCEGILEIFDMLATTSRFRKLQHKVEYLCVKAMIL 321
QY 346 RSIDHPGKLIAPDVLVDRDEGKCEGILEIFDMLATTSRFRKLQHKVEYLCVKAMIL 405
DB 322 LNSSMYPVLTATODADSSRKLAHLNAYTDALVWVIAKSGISSQOQSVRLANLMLSHV 381
QY 406 LNSSMYPVLTATODADSSRKLAHLNAYTDALVWVIAKSGISSQOQSVRLANLMLSHV 465
DB 382 RHISNKGMEHLISMCKKNVVPYDILLEMLANHTLRGYSISGESSSTEDSKNKSQ 441
QY 466 RHASNMKGMEHLISMCKKNVVPYDILLEMLANHTLRGYSISGESSSTEDSKNKSQ 525
DB 442 NLOSO 446
QY 526 NPOSO 530

```

```

RESULT 11
ID 055016 PRELIMINARY; PRT; 464 AA.
AC 055016;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE ESTROGEN RECEPTOR BETA2 DELTA3.
GN ERBETA.
OS RATRUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
RA BUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAUE-DAMLEY;
RA PETERSEN D.N., TKALCEVIC G.T., KOZA-TAYLOR P.H., TURI T.G.,
RA BROWN T.A.;
RL ENDOCRINOLOGY 0:0-0(1998).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: AF042061; G2801697;
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW ZINC-FINGER.
SQ SEQUENCE 464 AA: 51715 MW: 70679086 CRC32:

```

Query Match 60.1%; Score 2347; DB 11; Length 464;
 Best Local Similarity 78.1%; Pred. No. 0.00e+00;
 Matches 393; Conservative 34; Mismatches 19; Indels 57; Gaps 2;

```

DB 1 MITYSPAVMYVSGTNSNDGPGVROSTSPNVLMPTSGHLSPLATHGSSLLYAEPOKS 60
QY 46 MITYSPAVMYVSGTNSNDGPGVROSTSPNVLMPTSGHLSPLATHGSSLLYAEPOKS 105
DB 61 PWCBAFSLHTLPVNRKTLKRVSGNRCASVYTGPGSKRDHFCACVSDYASGYHGWMS 120
QY 106 PWCBAFSLHTLPVNRKTLKRVSGNRCASVYTGPGSKRDHFCACVSDYASGYHGWMS 165
DB 121 CEGCKAFKRSIQG-----SRRERC 141
QY 166 CEGCKAFKRSIQGHNDYICPATNQCTIDKNRRKSCQACRLKCYEYGMVCKGSRRCRG 225
DB 142 YRIVRORSSSEVYCHCSKAKRNGGAPRVKELLSTLSPQVLVTLLEAPPVLSRP 201

```

```

QY 226 YRIVRORSADEQLHCAGKAKRSRGHAPRVRELLDLALSPQVLVTLLEAPPVLSRP 285
DB 202 SMPTEASMMASLTKLADKELVHMIGMAKKIPGFVELSLDQVRLSCMVEVLMGLMW 261
QY 286 SAPTEASMMASLTKLADKELVHMISMAKKIPGFVELSLDQVRLSCMVEVLMGLMW 345
DB 262 RSIDHPGKLIAPDVLVDRDEGKCEGILEIFDMLATTSRFRKLQHKVEYLCVKAMIL 321
QY 346 RSIDHPGKLIAPDVLVDRDEGKCEGILEIFDMLATTSRFRKLQHKVEYLCVKAMIL 405
DB 382 RHISNKGMEHLISMCKKNVVPYDILLEMLANHTLRGYSISGESSSTEDSKNKSQ 441
QY 448 SGOQSVRLANLMLSHVHRISNKGMEHLISMCKKNVVPYDILLEMLANHTLRGYSIS 507
QY 508 TGSECSPAEDSKSGESSQNPQSO 530

```

```

RESULT 12
ID 013012 PRELIMINARY; PRT; 573 AA.
AC 013012;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE ESTROGEN RECEPTOR.
GN ANGUILLA JAPONICA (JAPANESE EEL).
OS EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC OSTEIFICHTHYES; ACTINOPTERYGII; ANGUILLIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LIVER;
RA TODO T.;
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-LIVER;
RA TODO T., ADACHI S., YAMAUCHI K.;
RL MOL. CELL. ENDOCRINOLOGY 119:37-45(1996).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: AB003356; D1020638;
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
DR PFAM: PF00104; hormone-rec.
DR PFAM: PF00105; zf-C4.
KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW ZINC-FINGER.
SQ SEQUENCE 573 AA: 63420 MW: A0B07849 CRC32:

```

Query Match 57.4%; Score 2240; DB 13; Length 573;
 Best Local Similarity 58.9%; Pred. No. 0.00e+00;
 Matches 310; Conservative 101; Mismatches 98; Indels 17; Gaps 13;

```

DB 33 PUYNVALALMESNAVIPSPTDSSHDYALTFYSPILISHGPA--VP--ESNAHQ 89
QY 13 PUYNVALALMESNAVIPSPTDSSHDYALTFYSPILISHGPA--VP--ESNAHQ 72
DB 90 SUSLFWFAHGHVSPDLALFQOPLVYRPAHSPMAEPKPLEHGOQTSKLAGKRA 149
QY 73 TISPNTLWPTPGH--LSPVLVYHROLSHLXAPKSPWCEAFSLHTLP--VNRKTLKRV 128
DB 150 ESEEGTSVGGCPAGKGDHMFCAVCHDVASGYHYGVSCGCKAFKRSIQGHNGYICPA 209
QY 129 SGNRCASVYTGPG--GSKRDHFCACVSDYASGYHYGVSCGCKAFKRSIQGHNDYICPA 187
DB 210 TNOCTIDKNRRKSCQACRLKCYEYGMVCKGSRRCRGYRIVRORSADE--QLH--CAG-- 269
QY 188 TNOCTIDKNRRKSCQACRLKCYEYGMVCKGSRRCRGYRIVRORSADE--QLH--CAG-- 243

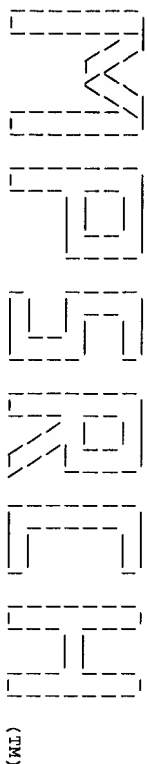
```


RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 96103562.
 RA YOUNG L.O., GORDIN J., GRAMER M., GAHR M., CREWS D.:
 RL J. STEROID BIOCHEM. MOL. BIOL. 55:261-269(1995).
 DR EMBL; S79923; G1195592; -.
 DR PFAM; PF00104; hormone_rec.
 DR PFAM; PF00105; zf-C4.
 FT NON_TER 1 1
 SO SEQUENCE 307 AA; 35202 MW; DC0871D8 CRC32;

Query Match	31.4%;	Score 1227;	DB 13;	Length 307;
Best Local Similarity	56.7%;	Pred. No. 1.08e-255;		
Matches 174;	Conservative	68;	Mismatches 53;	Indels 12;
				Gaps 10;

1	GHNDFMCAITNOCTIDKRRKSCACARJCYEYGMXKGGIRKXRRGRGLTKRROEEH	60
179	GHNDFICATNOCTIDKRRKSCACARJCYEYGMXKGGIRKXRRGRGLTKRROEEH	236
Dd	61 DNRNAGAIYERSPNMFSPMLTHHNKNKSPALSLADQIVASALLEAPPVYTSYEDPSR	120
Qy	237 EQLHCAGKA-KRSG--HAP-RV-RELLD-ALS-EEQVLTLLLEAPPHVLTIS-RPSA	287
Dd	121 PFSEASMTLLTINLADRELVMIMAKRVPGFYDLSHDQVHLECAWLETIMGLWRS	180
Qy	288 PFTEASNMSTIKLADKRLVMISMAKKIPGFVLSLFDQVRLLESCEVMIMGLMWRS	347
Dd	181 VEHQCKLFAFNLLIDNRQGCVCGEYEIFDMILATSSRRPMNVQGEFVCLKSTILLN	240
Qy	348 IDHFKLIFADVDLDRREGCVGELIEIFDMILATSSRRRELKLOHKEVLCYKAMILLN	407
Dd	241 SGITYFSLSTLSLEKHDIRHVRDLKLIIDTLHLMAKSGLSLQOQHRLAQLLILSHFR	300
Qy	408 SSMTPLVATGDA-DSRKLAHLNLNAVYDALVWVIANSGLSSQOQSRMLNLLMLSHVR	466
Dd	301 EMSNKGK 307	
Qy	467 HASNKGK 473	

Search completed: Mon Nov 2 16:36:13 1998
Job time : 232 secs.



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSRCH_PP protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Nov 2 16:36:30 1998; MasPar time 7.02 Seconds

Tabular output not generated. 534.685 Million cell updates/sec

Title: >US-08-906-365-2
Description: (1-530) From US08906365.pep
Perfect Score: 3904
Sequence: 1 MDIKNSPSSINSPPSYNCQ.....ECPSPEDSKSKESQNPQSQ 530

Scoring table: PAM 150
Gap 11

Searched: 77309 seqs, 7078906 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5_COMB 2:PCT9_COMB 3:backfiles1

Statistics: Mean 33.550; Variance 151.595; scale 0.221

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	1648	42.2	410	3	5223606-5 Patent No. 5223606.	5.61e-141
2	660	16.9	363	3	5223606-6 Patent No. 5223606.	1.47e-48
3	609	15.6	356	3	5223606-7 Patent No. 5223606.	7.04e-44
4	577	14.8	795	1	US-07-716-Sequence 5, Applicatio	5.94e-41
5	533	13.6	533	1	US-07-952-Sequence 2, Applicatio	6.07e-37
6	539	13.6	446	1	US-07-952-Sequence 4, Applicatio	1.40e-36
7	514	13.2	467	1	US-08-336-Sequence 4, Applicatio	3.23e-35
8	514	13.2	467	2	PCT-US91-0-Sequence 4, Applicatio	3.23e-35
9	513	13.1	462	1	US-08-336-Sequence 2, Applicatio	3.98e-35
10	513	13.1	462	2	PCT-US91-0-Sequence 6, Applicatio	1.11e-32
11	466	12.4	463	1	US-08-336-Sequence 6, Applicatio	1.11e-32
12	466	12.4	463	2	PCT-US91-0-Sequence 6, Applicatio	1.11e-32
13	441	11.3	454	3	5260432-2 Patent No. 5260432.	1.27e-28
14	439	11.2	368	3	5223606-3 Patent No. 5223606.	1.91e-28
15	439	11.2	448	3	5223606-2 Patent No. 5223606.	1.91e-28
16	438	11.0	462	2	PCT-US92-0-Sequence 4, Applicatio	1.86e-27
17	438	11.0	462	2	PCT-US92-0-Sequence 2, Applicatio	9.69e-27
18	430	10.8	440	1	US-08-333-Sequence 8, Applicatio	9.69e-27
19	420	10.8	440	1	US-08-694-Sequence 8, Applicatio	9.69e-27
20	420	10.8	440	1	US-08-463-Sequence 8, Applicatio	9.69e-27
21	420	10.8	447	1	US-08-373-Sequence 1, Applicatio	9.69e-27
22	413	10.6	462	3	5171671-2 Patent No. 5171671.	4.10e-26
23	378	9.7	472	1	US-08-496-Sequence 2, Applicatio	5.40e-23

24	365	9.3	348	1	US-07-843-Sequence 10, Applicati	7.68e-22
25	365	9.3	348	1	US-08-458-Sequence 10, Applicati	7.68e-22
26	365	9.3	348	1	US-08-459-Sequence 10, Applicati	7.68e-22
27	358	9.2	414	1	US-08-255-Sequence 9, Applicatio	3.20e-21
28	361	9.2	469	2	PCT-US95-1-Sequence 9, Applicatio	1.74e-21
29	357	9.1	89	1	US-08-087-Sequence 2, Applicatio	3.92e-21
30	357	9.1	89	1	US-08-561-Sequence 3, Applicatio	8.85e-21
31	353	9.0	484	2	PCT-US95-1-Sequence 1, Applicatio	2.00e-20
32	349	8.9	66	3	5217867-1 Patent No. 5217867.	1.15e-18
33	329	8.4	451	2	PCT-US95-1-Sequence 2, Applicatio	1.92e-17
34	315	8.1	70	1	US-07-737-Sequence 3, Applicatio	9.64e-17
35	307	7.9	443	1	US-08-342-Sequence 4, Applicatio	9.64e-17
36	307	7.9	446	2	PCT-US95-1-Sequence 3, Applicatio	5.28e-17
37	310	7.9	461	1	US-08-330-Sequence 2, Applicatio	5.28e-17
38	310	7.9	461	1	US-08-330-Sequence 2, Applicatio	5.28e-17
39	310	7.9	461	2	PCT-US95-1-Sequence 2, Applicatio	5.28e-17
40	310	7.9	461	2	PCT-US95-1-Sequence 2, Applicatio	5.28e-17
41	309	7.9	548	1	US-08-333-Sequence 2, Applicatio	6.45e-17
42	309	7.9	548	1	US-08-694-Sequence 2, Applicatio	6.45e-17
43	309	7.9	548	1	US-08-463-Sequence 2, Applicatio	6.45e-17
44	304	7.8	558	1	US-08-463-Sequence 12, Applicati	1.76e-16
45	304	7.8	558	1	US-08-333-Sequence 12, Applicati	1.76e-16

ALIGNMENTS

RESULT	ID	1	STANDARD;	PRT;	444	AA.
XX	5223606-5	xxxxxx				
XX	01-JAN-1900					
DE	Patent No. 5223606.					
CC	Patent No. 5223606					
CC	APPLICANT: BLAUDIN DE THE, HUGHES; MARCHIO, AGNES; TIOLLATS,					
CC	PIERRE-DEJEAN, ANNE					
CC	TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED					
CC	PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA					
CC	NUMBER OF SEQUENCES: 11					
CC	CURRENT APPLICATION NUMBER: US-07/134,130					
CC	APPLICATION NUMBER: US-07/134,130					
CC	FILING DATE: 17-DEC-1987					
CC	PRIOR APPLICATION DATA:					
CC	SEQ ID NO:5:					
CC	LENGTH: 410					
CC	SEQUENCE 444 AA; 50020 MW; 1110634 CN;					
DB	Query Match	42.2%	Score 1648;	DB 3;	Length 410;	
DB	Best Local Similarity	56.9%	Pred. No. 5.61e-141;			
DB	Matches	224;	Conservative 92;	Mismatches 64;	Indels 14;	Gaps 11;
DB	1 CAVCNDYASGYHYGWSCEGCAKAFKRSIOGHNDWCPATNCTIDKNRKSQCACRLK 60					
DB	149 CAVCSYASGYHYGWSCEGCAKAFKRSIOGHNDWCPATNCTIDKNRKSQCACRLK 208					
DB	61 CEVGVMMKGIKRRGRGRLKHKRRDDGEGRGVGSAGDMAANLWPSLMIRSKKN 120					
DB	209 CEVGVMMKGIKRRGRGRLKHKRRDDGEGRGVGSAGDMAANLWPSLMIRSKKN 259					
DB	121 STALSTADQWYVALIDAPILYSEYDPRPSEASMGGLTNLADRELVMINNAKRYV 180					
DB	260 -LD-ALSPDVLVTLLEAPPHVLISRPSPAPTEASMMMSLTKLADKELVHMIWNAKKIP 317					
DB	181 GPVDTLHDOVHLLECAWELIIMIGLVWRSMEHPVKLFAPLVLLDRNGCKVEGVETP 240					
DB	318 GVEVLSLFDQVRLSCWMEVLMGMWRSIDHPGLIFAPLVLDROGCKVEGVETP 377					
DB	241 DMLATSSRRFMMNLOGEFFVCLKSIILNSGVYFLSTLSLEKDIHVLDKITTY 300					
DB	378 DMLATSSRRFMMNLOGEFFVCLKSIILNSGVYFLSTLSLEKDIHVLDKITTY 436					

QY 327 QVRLLESCWMEVLMGMIMRSIDHPGKLIFAPDLYLDRDEGKCVGEILIEFDMLLAT-TS 385

DB 356 KMRDMQMDKTELGLRAIVLPNDPSKGLSNPA-EVEA--L-R--EKVYASLE-AYCHK 407

QY 386 RFRELKQHEYLCKVAMILLNSMYPVLTATODDSSRKLAHLINAVTALVWVIKSG 445

DB 408 YP-EDPG-RFARLLRLPALRSIGKLEHLEFFFKLIGDTPIDTFLMEXEA 457

QY 446 ISSQOOSMRLANMLMLLSHVHRSNKGMEHLNMRCKNVVYDLEMLNA 497

RESULT 10
ID PCT-US91-00399-2 STANDARD; PRT; 462 AA.
XX xxxxxx

DE Sequence 2, Application PC/TUS9100399

CC GENERAL INFORMATION:
CC APPLICANT: Mangelsdorf, Dr., David J.
CC APPLICANT: Evans Dr., Ronald M.
CC TITLE OF INVENTION: RETINOID RECEPTOR COMPOSITIONS AND
CC NUMBER OF SEQUENCES: 7
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: FITCH, EVEN, TABIN & FLANNERY
CC STREET: 135 South LaSalle Street, Suite 900
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60603

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentln Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US91/00399
CC FILING DATE: 19910122
CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/478,071
CC FILING DATE: 09-FEB-1990

CC ATTORNEY/AGENT INFORMATION:
CC NAME: Reiter, Mr., Stephen E.
CC REGISTRATION NUMBER: 31192
CC REFERENCE/DOCKET NUMBER: 50852

CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (619) 552-1311
CC TELEFAX: (619) 552-0095
CC TELETYPE: 20 6566 PATLAW CGO

CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 462 amino acids
CC TYPE: AMINO ACID
CC TOPOLOGY: linear

CC MOLECULE TYPE: protein
CC SEQUENCE 462 AA; 50811 MW; 1063517 CN;

Query Match 13.1%; Score 513; DB 2; Length 462;
Best Local Similarity 33.0%; Pred. No. 3.98e-35;
Matches 116; Conservative 89; Mismatches 119; Indels 28; Gaps 21;

DB 133 HICATIGDSSGKHGVCCEGKGFEKRTVRKDLTYTGRDNKDCILIDRKQRQCYCY 192

QY 147 HFCVAVSDYASGHHGVCCEGKAFKRSIGGHNHYICPATNQCTIDNNRKSQACPL 206

DB 193 ORCLAMGRKRAVQEEH-O-R-GK-DRNENE-VESTSSAN--EDMPVER-ILEAVLAVE 243

QY 207 RRCYGVGWKVCGRRRRCGYRLVRRORSADQLHCGKAKRSGHAPRYVELLDALSPF 266

DB 244 PKEITYVEANMG-L-N-PSSP-NDP--VTNICOADKOLFVEMAKRIPHFSELPDD 296

QY 267 QVLTLLLEAPPHVILSRSPAFTEASMMSLTXLADKFLVHMIMAKRIIPFVELSLFD 326

DB 297 QVILLRAGNNELLISFSRSTAVKDGILLATGALVHRSNASHA-GVGAIFRVLTELV 355

QY 327 QVRLLESCWMEVLMGMIMRSIDHPGKLIFAPDLYLDRDEGKCVGEILIEFDMLLAT-TS 385

DB 356 KMRDMQMDKTELGLRAIVLPNDPSKGLSNPA-EVEA--L-R--EKVYASLE-AYCHK 407

QY 386 RFRELKQHEYLCKVAMILLNSMYPVLTATODDSSRKLAHLINAVTALVWVIKSG 445

DB 408 YP-EDPG-RFARLLRLPALRSIGKLEHLEFFFKLIGDTPIDTFLMEXEA 457

QY 446 ISSQOOSMRLANMLMLLSHVHRSNKGMEHLNMRCKNVVYDLEMLNA 497

RESULT 11
ID US-08-336-408B-6 STANDARD; PRT; 463 AA.
XX xxxxxx

DE Sequence 6, Application US/08336408B

CC Patent No. 5723329

CC GENERAL INFORMATION:
CC APPLICANT: EVANS, RONALD M.
CC APPLICANT: MANGELSDORF, DAVID J.
CC TITLE OF INVENTION: RETINOID RECEPTOR COMPOSITIONS AND
CC NUMBER OF SEQUENCES: 7
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
CC STREET: 444 SOUTH FLOWER STREET, SUITE 2000
CC CITY: LOS ANGELES
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 90071

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patentln Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/336,408B
CC FILING DATE: 08-NOV-1994

CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/933,453
CC FILING DATE: 21-AUG-1992

CC APPLICATION NUMBER: WO PCT/US91/00399
CC FILING DATE: 22-JAN-1991

CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/478,071
CC FILING DATE: 09-FEB-1990

CC ATTORNEY/AGENT INFORMATION:
CC NAME: REITER, STEPHEN E.
CC REGISTRATION NUMBER: 31,192
CC REFERENCE/DOCKET NUMBER: P41 9851

CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619-546-1995
CC TELEFAX: 619-546-9392

CC INFORMATION FOR SEQ ID NO: 6:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 463 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear

CC MOLECULE TYPE: protein

QY 149 CAVCSDYASGYHYGWSCEGCAFFKRSIQGHNDYICPATNOCTIKNRKRCQACRLK 208
DB 150 CFEVGSKEAVRND-R-N-K-KKKEVEEGSPDSYEI:SPOLELITRYSAHOETF-PS-L 204
QY 209 CFEVGVKCGSRRERCGYRLVRQRQSADEQLHCAGKAKRSGHAPRYRELLDALSPEDL 268
DB 205 C-QL-GKYTTNSSADHR-VOL-DGLMDKFSELATKCIITVEFAKRLPGFTGLIADQI 260
QY 269 VLTLLAEPPHVLISRPASAPFTESAMMSLTKLADKELVHMISWAKKIPGFVLSLFDQY 328
DB 261 TLKAC-LDILIRICTRTPEODTMTFSDGLINRTO-MHNAGFPLDVFAPAGOL 318
QY 329 RLL-ESCMVEVLMGLMMSIDHPGKLI:FPADVLDRDESKCYEGILEIFDMILATTSRF 387
DB 319 LPLEMDTE---TG--LL-SALICIGDRMDLEEPKVDKLOEPLLEAL-RLYARRRRP 370
QY 388 RELKLOHKEYLCYKAMILLNMSWPLVTATQDADSSKRLAHNLNAVYDALVWYIAKSGIS 447
DB 371 SQP-YM-FPRLMKITDRLGISTGAEARATLTMELIPGMPPLIREMLE 417
QY 448 SQOOSMRLANLMLSHVRHASNKGMEHLNMCKKNVVPYDILLEMLN 496

RESULT 14
ID 5223606-3 STANDARD; PRT; 398 AA.
XX
AC xxxxxx
XX
DT 01-JAN-1900
DE Patent No. 5223606.
XX
XX
CC Patent No. 5223606
CC APPLICANT: BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOILLAIS,
CC PIERRE;DEJEAN, ANNE
CC TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC NUMBER OF SEQUENCES: 11
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/134,130
CC FILING DATE: 17-DEC-1987
CC PRIOR APPLICATION DATA:
CC SEQ ID NO:3:
CC LENGTH: 368
CC
SQ SEQUENCE 398 AA; 44842 MW; 907111 CN;

Query Match 11.2%; Score 439; DB 3; Length 368;
Best Local Similarity 26.9%; Pred. No.1,91e-28;
Matches 94; Conservative 97; Mismatches 136; Indels 22; Gaps 18;

DB 1 CFWCQDKSSGYHYGWSACEGCKGFFRSIOKNMITYCHRDKNVCVINKYTRNRCQYCRLOK 60
QY 149 CAVCSDYASGYHYGWSCEGCAFFKRSIQGHNDYICPATNOCTIKNRKRCQACRLK 208
DB 61 CFEVGSKEAVRND-R-N-K-KKKEVEEGSPDSYEI:SPOLELITRYSAHOETF-PS-L 115
QY 209 CFEVGVKCGSRRERCGYRLVRQRQSADEQLHCAGKAKRSGHAPRYRELLDALSPEDL 268
DB 116 C-QL-GKYTTNSSADHR-VOL-DGLMDKFSELATKCIITVEFAKRLPGFTGLIADQI 171
QY 269 VLTLLAEPPHVLISRPASAPFTESAMMSLTKLADKELVHMISWAKKIPGFVLSLFDQY 328
DB 172 TLKAC-LDILIRICTRTPEODTMTFSDGLINRTO-MHNAGFPLDVFAPAGOL 229
QY 329 RLL-ESCMVEVLMGLMMSIDHPGKLI:FPADVLDRDESKCYEGILEIFDMILATTSRF 387
DB 230 LPLEMDTE---TG--LL-SALICIGDRMDLEEPKVDKLOEPLLEAL-RLYARRRRP 281
QY 388 RELKLOHKEYLCYKAMILLNMSWPLVTATQDADSSKRLAHNLNAVYDALVWYIAKSGIS 447
DB 282 SKPH-M-FPKILMKITDRLSISAKGAEVITLKMELIPGSMPLIOEME 328
QY 448 SQOOSMRLANLMLSHVRHASNKGMEHLNMCKKNVVPYDILLEMLN 496

RESULT 15
ID 5223606-2 STANDARD; PRT; 485 AA.
XX
AC xxxxxx
XX
DT 01-JAN-1900
DE Patent No. 5223606.
XX
XX
CC Patent No. 5223606
CC APPLICANT: BLAUDIN DE THE, HUGHES;MARCHIO, AGNES;TIOILLAIS,
CC PIERRE;DEJEAN, ANNE
CC TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
CC PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA
CC NUMBER OF SEQUENCES: 11
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/134,130
CC FILING DATE: 17-DEC-1987
CC PRIOR APPLICATION DATA:
CC SEQ ID NO:2:
CC LENGTH: 448
CC
SQ SEQUENCE 485 AA; 54437 MW; 1345912 CN;

Query Match 11.2%; Score 439; DB 3; Length 448;
Best Local Similarity 26.9%; Pred. No.1,91e-28;
Matches 94; Conservative 97; Mismatches 136; Indels 22; Gaps 18;

DB 81 CFWCQDKSSGYHYGWSACEGCKGFFRSIOKNMITYCHRDKNVCVINKYTRNRCQYCRLOK 140
QY 149 CAVCSDYASGYHYGWSCEGCAFFKRSIQGHNDYICPATNOCTIKNRKRCQACRLK 208
DB 141 CFEVGSKEAVRND-R-N-K-KKKEVEEGSPDSYEI:SPOLELITRYSAHOETF-PS-L 195
QY 209 CFEVGVKCGSRRERCGYRLVRQRQSADEQLHCAGKAKRSGHAPRYRELLDALSPEDL 268
DB 196 C-QL-GKYTTNSSADHR-VOL-DGLMDKFSELATKCIITVEFAKRLPGFTGLIADQI 251
QY 269 VLTLLAEPPHVLISRPASAPFTESAMMSLTKLADKELVHMISWAKKIPGFVLSLFDQY 328
DB 252 TLKAC-LDILIRICTRTPEODTMTFSDGLINRTO-MHNAGFPLDVFAPAGOL 309
QY 329 RLL-ESCMVEVLMGLMMSIDHPGKLI:FPADVLDRDESKCYEGILEIFDMILATTSRF 387
DB 310 LPLEMDTE---TG--LL-SALICIGDRMDLEEPKVDKLOEPLLEAL-RLYARRRRP 361
QY 388 RELKLOHKEYLCYKAMILLNMSWPLVTATQDADSSKRLAHNLNAVYDALVWYIAKSGIS 447
DB 362 SKPH-M-FPKILMKITDRLSISAKGAEVITLKMELIPGSMPLIOEME 408
QY 448 SQOOSMRLANLMLSHVRHASNKGMEHLNMCKKNVVPYDILLEMLN 496

Search completed: Mon Nov 2 16:37:14 1998
Job time : 44 secs.

GAATTCGCGACGAG 3' -3' adaptor sequence: 5'
 CTCGAGTTTTTTTTTTTTTTTTTTT 3',"
 /db_xref="GDB:5589314"
 /db_xref="taxon:9606"
 /clone="648794"
 /clone_lib="Stratagene hnt neuron (#937233)"
 /dev_stage="hnt neurons"
 /lab_host="SOLR (kanamycin resistant)"
 BASE COUNT 192 a 74 c 112 g 146 t 7 others
 ORIGIN

Query Match 10.3%; Score 173; DB 15; Length 531;
 Best Local Similarity 99.4%; Pred. No. 0.00e+00;
 Matches 174; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 27 CGTGTCCAGGGGTAGATGATGATGCTGAGTGTAGAGAGAAGATTAAGCTAG 86
 Cp 175 cgtgcctccagggtgaagatgagatgactgcagcttgtagagagaagaatgaagctag 116
 Db 87 ATGCTGAGTTTTTATATCCATGCTTGTAGATACAGCTGAGAAAACCTTGCAAGAAG 146
 Cp 115 atgtgagttttttatatacatgcttgtagatgacgtcagctgagaaaacaccttgcaagaag 56
 Db 147 AGGCACAAAGGTCATATATATGTTCTCAAGATTCCTGGGCAAGTATATGCTG 201
 Cp 55 aggcacaaaggtcatataatgcttccaagatcgtggtgcaagataatgctg 1

RESULT 2 AA744644 466 bp mRNA EST 07-FEB-1998
 LOCUS n79f02.s1 NCI-CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1284507
 DEFINITION similar to gb:MI2674 ESTROGEN RECEPTOR (HUMAN);, mRNA sequence.
 ACCESSION AA744644
 NID 92783408
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 466)
 NCBI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Louis W. Staudt, M.D., Ph.D., David Allman,
 Ph.D., Gerald Marti, M.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/dbp/image/image.html
 Insert length: 1369 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 430.
 Location/Qualifiers
 1. 466
 /organism="Homo sapiens"
 /note="vector: p773D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from human tonsillar cells enriched for
 germinal center B cells by flow sorting (CD20+, IgD-),
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
 (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
 primed with a Not I - oligo(dT) primer
 [5'-TGTTCACATCTGAGTGGAGGCGCCGCTCATTTTCTTTTCTTTT-
 3']. Double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified p773 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."
 /db_xref="taxon:9606"
 /clone="IMAGE:1284507"
 /clone_lib="NCI-CGAP_GCB1"
 /tissue_type="germinal center B cell"
 /lab_host="DH10B"
 BASE COUNT 113 a 134 c 130 g 88 t 1 others
 ORIGIN

Query Match 4.5%; Score 76; DB 13; Length 466;
 Best Local Similarity 82.8%; Pred. No. 2.42e-105;
 Matches 96; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
 Db 351 GGAGCTGCTCTGTGAGGCGTCAAGGCTCTTCAAGAGAAGATTCAGACATAAC 410
 Qy 577 ggaagctgctgtgagagagatgagatgagccttttaagaagaagatcaagacataat 636
 Db 411 GACTATATGTGTCCAGCCACCAACCAAGTGCACCATTTGATTAACAGAGAGAAG 466
 Qy 637 gattatattgtccagctacacataatcagtgtaacatcgataaataaacgcgcgaag 692

RESULT 3 AA754459 252 bp mRNA EST 20-JAN-1998
 LOCUS 97SN1787 Rice Immature Seed Lambda ZAPIT cDNA Library Oryza sativa
 DEFINITION cDNA clone 97SN1787, mRNA sequence.
 ACCESSION AA754459
 NID 92801165
 KEYWORDS EST.
 SOURCE rice.
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Charophyta/Embryophyta group;
 Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
 Poales; Poaceae; Oryza.
 REFERENCE 1 (bases 1 to 252)
 Nalm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
 Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
 Lee,M.C. and Eun,M.Y.
 Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
 Unpublished (1998)
 JOURNAL
 TITLE
 COMMENT

Contact: Eun M.Y.
 Department of Cytogenetics
 National Inst. of Agri. Sci. and Tech, RDA
 Suwon, Kyunggi-do, Korea
 Tel: 82 331 290 0301
 Fax: 82 331 290 0307
 Email: myeunsun20.asl.re.kr
 Submitted by Baek Hie Nalm, Dept of Biological Science, Myongji
 University, Yongin, Korea. 449-728 bhnalm@ioserver.myongji.ac.kr
 Seq primer: M13 Reverse Primer.
 Location/Qualifiers
 1. 252
 /organism="Oryza sativa"
 /cultivar="Milyang23"
 /note="vector: p Bluescript SK(+); Site_1: EcoRI; Site_2:
 XhoI; Directional cDNA library inserted into lambda ZAPIT
 vector at 5' end with EcoRI and 3' end with Xho I site."
 /db_xref="taxon:4530"
 /clone="97SN1787"
 /clone_lib="Rice Immature Seed Lambda ZAPIT cDNA library"
 /tissue_type="Immature Seed"
 /dev_stage="5 days after pollination"
 /lab_host="E. coli SOLR"

BASE COUNT 5 a 21 c 12 g 35 t 179 others
 ORIGIN
 Query Match 4.0%; Score 68; DB 13; Length 252;
 Best Local Similarity 11.8%; Pred. No. 4.00e-88;
 Matches 27; Conservative 123; Mismatches 75; Indels 4; Gaps 3;

OY	983	ccgcagcaaatgtgccgacaagaagtctgtacatcatgatcagctggcgccaagaagatttc	1042
Db	127	BMBBGTGCTTTMCCWBHYNKTCIASCAGHHSTNTNDVKSSTTWGYTBSYDXSMHGWCGB	186
OY	1043	cgcgcttgttgaagctcagccctgtctgcaccagaagtcgagctcttgtagagactgttgatcgy	1102
Db	187	BVKHTTKVSTTRATRS-TCVRYKCYVMNTEKKYVKKYHVABGCBBHDSCCKIMMTNK	244
OY	1103	aggctctaatagtatggatggatgtgcctcatcatgaccaccgcggaagctcatcttt	1161
RESULT	5		
LOCUS	AA517920	534 bp	mRNA EST 14-JUL-1997
DEFINITION	vhp5e40.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone 902038 5' similar to gb:M4476 Mouse retinoic acid receptor gamma (MOUSE); mRNA sequence.		
ACCESSION	AA517920		
NID	G2257444		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 534) Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Gaisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Weising,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and Westrop,R.		
TITLE	The WashU-HMI Mouse EST Project		
JOURNAL	Unpublished (1996)		
COMMENT	Contact: Marra M/Mouse EST Project WashU-HMI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:522722 Seq primer: -28ml3 rev2 ET from Amersham High quality sequence stop: 393. Location/Qualifiers 1. 534 /organism="Mus musculus" /strain="C3H" /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site.1: EcoRI; Site.2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTAGCAATCTGAAGCGGAGGGCGGCCCTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [AATTGGATCCTTG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed by Bob Barstead. The C2C12 cell line (available from ATCC, catalog # CRL-1772) differentiates rapidly, forming contractile myotubes and producing characteristic muscle proteins." /db_xref="taxon:10090" /clone="902038" /clone_1lb="Barstead mouse myotubes MPLRB5" /cell_line="C2C12" /lab_host="DH10B" <1..-534		
BASE COUNT	128 a	153 c	143 g 110 t
ORIGIN	mRNA		
Query Match	3 6%	Score 61;	DB 7; Length 534;
Best Local Similarity	66.7%;	Pred. No. 2.19e-73;	
Matches	132; Conservative	0; Mismatches 65;	Indels 1; Gaps 1;

Db	315	TCGCTTGGATCGAAGAACAGCTTCTGGCTACACATGAGGGGTGACGTCCTGTGAAGGC	374
Oy	538	tgcgtctgtcgaagsgattacgcacatccggatatacctaataatgagcttgcgtgtaaga	597
Db	375	TGCAAGGGCTTCTTGACAGCAGCATTCAGAGAAACATGTTGTATACATGTACCGTGACA	434
Oy	538	tgtaaagccctttttaaagaagcattcaagaagacataatgattatatttgcacgtaca	657
Db	435	AAA-ACGTATCATCAACAAGGTACACCAAGAAATCGATCGCACTACTCGACGGCTACAAAG	493
Oy	658	aatcgtgtgaacatcgtataaanaacccgcgcgaagcgtccagcgtccgcgacttcgaag	717
Db	494	TGTTTCGAAGTGGCGATG	511
Oy	718	tgttacgaagtgygaatg	735
RESULT	6	AA893618	384 bp mRNA EST 03-APR-1998
LOCUS		ESL197421	Normalized rat placenta, Bento Soares Rattus sp. cDNA
DEFINITION		clone RP.LAC64 3' end, mRNA sequence.	
ACCESSION		AA893618	
NID		g3020497	
KEYWORDS		EST.	
SOURCE		Rattus sp.	
ORGANISM		Rattus sp.	
REFERENCE		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
AUTHORS		1 (bases 1 to 384) Lee,N.H., Glodex,A., Chandra,I., Mason,T.M., Quackenbush,J., Kerlavage,A.R. and Adams,M.D.	
TITLE		Rat Genome Project: Generation of a Rat EST (RESt) Catalog & Rat Gene Index	
JOURNAL		Unpublished (1998)	
COMMENT		Contact: Lee, NH ATCC The Institute for Genomic Research 9712, Medical Center Drive, Rockville, MD 20850, USA Tel: (301)-838-3529 Fax: (301)-838-0208 Email: nhlee@tigr.org Seq primer: M13-21.	
FEATURES		Location/Qualifiers	
Source		1..384 /organism="Rattus sp." /note="Organ: placenta; Vector: pT7mPac; Site_1: ECoRI Site_2: NotI" /db_xref="taxon:10118" /clone="RPLC64" /clone_id="Normalized rat placenta, Bento Soares"	
BASE COUNT		87 a 102 c 94 g 101 t	
ORIGIN			
Query Match		3.3%; Score 56; DB 16; Length 384;	
Best Local Similarity		64.0%; Pred. No. 4.37e-63;	
Matches 128; Conservative		0; Mismatches 72; Indels 0; Gaps 0;	
Db	15	CATTCACGCCCTGGAAGACATTTCCGATACGGCAGCATGGGAGTTTTCCTCGAATTT	74
Cp	735	cattccacccttcgtaaacctccgaagtcggaagcctgtgacgtcttgcgcggttttc	676
Db	75	ATCAATGATGCAATCGTTTCTTCCAGACAAGGTAATTTGCTGCTCCACTGCTCT	134
Cp	675	atcgatgtacacgatttgcgtagcttgacaaataatacatatagcttcgtgaatgcttc	616
Db	135	TTTAAAGAAATCTTTGGACGCTTCCACATGTGACGACCCCGTAAATGACCTGAGCTTC	194
Cp	615	tttaaaaaggccttaccttccttcacacgaccagacatcctatagatataccgatgcta	556
Db	195	ATCGAGCACCCAGGCGA	214
Cp	555	atcgctcgacagcagcgaga	536

RESULT	7	AA851319	384 bp	mRNA	EST	30-APR-1998
LOCUS		EST194087	Normalized	rat placenta,	Bento Soares	Rattus sp. cDNA
DEFINITION		clone RP1AF19 3' end,	mRNA sequence.			
ACCESSION		AA851319				
NID		92938859				
KEYWORDS		EST.				
SOURCE		Rattus sp.				
ORGANISM		Rattus sp.				
		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;				
		Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
REFERENCE		1 (bases 1 to 384)				
AUTHORS		Lee,N.H., Glodok,A., Chandra,I., Mason,T.M., Quackenbush,J.,				
		Kerlavage,A.R. and Adams,M.D.				
TITLE		Rat Genome Project: Generation of a Rat EST (RESt) Catalog & Rat				
JOURNAL		Gene Index				
COMMENT		Unpublished (1998)				
		Contact: Lee, NH				
		ATCC				
		The Institute for Genomic Research				
		9712, Medical Center Drive, Rockville, MD 20850, USA				
		Tel: (301)-838-3529				
		Fax: (301)-838-0208				
		Email: nhlee@tigr.org				
		Seq primer: M13-21.				
FEATURES		Location/Qualifiers				
source		1..384				
		/organism="Rattus sp."				
		/note="Organ: Placenta; Vector: pT7T3pac; Site_1: EcoRI;				
		Site_2: NotI"				
		/db_xref="ATCC (inhost):2011237"				
		/db_xref="taxon:10118"				
		/clone="RP1AF19"				
		/clone_1lb="Normalized rat placenta, Bento Soares"				
BASE COUNT		87 a	102 c	94 g	101 t	
ORIGIN						
		Query Match	3.3%;	Score 56;	DB 16;	Length 384;
		Best Local Similarity 64.0%;		Pred. No. 4.37e-63;		
		Matches 128;	Conservative 0;	Mismatches 72;	Indels 0;	Gaps 0;
Db	15	CATTCCAGCCTGAGACATTTCCGATAGCGGCACTGCTGGCAGTTTTCCTCGAATTTT	74			
Cp	735	cattccagccttgagacattccgaagtcgagcgcttgccgctcttgccgcggtttt	676			
Db	75	ATCATATGATGATCGTTCGTTCTTCCACGACAAAGTATGTGCTGTCCTCCACATGCTCT	134			
Cp	675	atcgatgtcacctgatttgtagtcggaacaataataatcattatgctcttgaaagctctc	616			
Db	135	TTTAAGATATCTTTGCAGGCTTCACATGTCAGACCCCGTATGACATCGTAAGATTTC	194			
Cp	615	tttaaaagagccttcatccttcacacagcagcagactcatatgatatacccgatgcgta	556			
Db	195	ATCGGAGCACACACGACGACGA	214			
Cp	555	atcgctgcagacagcgacga	536			
RESULT	8	AA754458	247 bp	mRNA	EST	20-JAN-1998
LOCUS		97SN1784	Rice	Immature Seed	Lambda ZAPII	cDNA Library
DEFINITION		cDNA clone 97SN1784,	mRNA sequence.			
ACCESSION		AA754458				
NID		92801164				
KEYWORDS		EST.				
SOURCE		rice.				
ORGANISM		Oryza sativa				
		Eukaryota; Viridiplantae; Charophyta/Embryophyta group;				
		Monocotyledonae; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;				
		Poales; Poaceae; Oryza.				

double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). The retinas were obtained from a 55 year old Caucasian male and total cellular poly(A)⁺ RNA was extracted 6 hrs after their removal. The retina RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the University of Toronto. Library constructed by Bento Soares and M.Fatima Bonaldo.

ORGANISM

Homo sapiens
Eukaryotes; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 403)

REFERENCE
AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE
JOURNAL
COMMENT

The WashU-Merck EST Project
Unpublished (1995)

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
High quality sequence stops: 330
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES

source 1..403
/organism="Homo sapiens"
/clone="221197"

BASE COUNT 73 a 113 c 118 g 97 t 2 others
ORIGIN

Query Match 3.1%; Score 52; DB 21; Length 403;

Best Local Similarity 66.7%; Pred. No. 5.21e-55; Matches 104; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Db 112 CTTGACATCCCATCTTTGAGCATTTTCATGAGCGGCGAGCGCTGCGGCT 171
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 741 ctccaccattccacactctgtaacactccgaagtcgacgctgcagctctgcgcg 682
Db 172 CCGTTGGATCTCGCACTCGTTGGTGGCGGCGAGCTGACTCAATGTTCCCTGGAT 231
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 681 gttttatcagatgtacactgattgtagctgacaaataataatcatatgctcctgaat 622
Db 232 AGTCCCTTGAAGAGGCGCTTGCAAGCCTGCGAGGA 267
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 621 gctcttttaaaaaagccttaccatccttcacagca 566

RESULT 11 H82542 454 bp mRNA EST 09-NOV-1995
LOCUS ys69h11.r1 Homo sapiens cDNA clone 220101.5' similar to
DEFINITION gb:351417_cds1 STEROID HORMONE RECEPTOR ERR2 (HUMAN);
ACCESSION H82542
NID 91060631
KEYWORDS EST.
SOURCE human clone-220101 primer-M13R1 library-Soares retina N24HR
vector-pT73D (Pharmacia) with a modified polylinker host-DH10B
(ampicillin resistant) Rsite1-Not I Rsite2-Eco RI 1st strand cDNA
was primed with a Not I - Oligo(dT) primer
[5'-TGTACCAATCTGAGCGGCGGCGCTTTTCTTTTCTTTTCTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT73 vector (Pharmacia). The retinas were

obtained from a 55 year old Caucasian male and total cellular poly(A)⁺ RNA was extracted 6 hrs after their removal. The retina RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the University of Toronto. Library constructed by Bento Soares and M.Fatima Bonaldo.

ORGANISM

Homo sapiens
Eukaryotes; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 454)

REFERENCE
AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE
JOURNAL
COMMENT

The WashU-Merck EST Project
Unpublished (1995)

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
High quality sequence stops: 345
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES

source 1..454
/organism="Homo sapiens"
/clone="220101"

BASE COUNT 82 a 118 c 136 g 114 t 4 others
ORIGIN

Query Match 3.1%; Score 53; DB 21; Length 454;

Best Local Similarity 65.7%; Pred. No. 5.16e-57; Matches 119; Conservative 0; Mismatches 61; Indels 1; Gaps 1;

Db 137 CTTGACATCCCATCTTTGAGCATTTTCATGAGCGGCGAGCGCTGCGGCT 196
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 741 ctccaccattccacactctgtaacactccgaagtcgacgctgcagctctgcgcg 682
Db 197 CCGTTGGATCTCGCACTCGTTGGTGGCGGCGAGCTGACTCAATGTTCCCTGGAT 256
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 681 gttttatcagatgtacactgattgtagctgacaaataataatcatatgctcctgaat 622
Db 257 AGTCCCTTGAAGAGGCGCTTGCAAGCCTGCGAGGAGCAAGCCTNAGTGTAGCAG 316
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 621 gctcttttaaaaaagccttaccatccttcacacagac-cagactccatagatgatacccg 563
Db 317 A 317
Cp 562 a 562

RESULT 12 AA411285 462 bp mRNA EST 17-MAY-1997
LOCUS zv26g03.r1 Soares NHMPA S1 Homo sapiens cDNA clone 754804.5'
DEFINITION similar to gb:M29960 STEROID RECEPTOR TR2 (HUMAN);, mRNA sequence.
ACCESSION AA411285
NID 92068936
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 462)
REFERENCE Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,

TITLE
JOURNAL
COMMENT

Kucaba, T., Lacy, M., Le, N., Lennon, G., Maris, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
MashU-Merck EST Project 1997
Unpublished (1997)

OY 121 RETLKRKYSGRKASPVTPGSGSKRDHAFCAVCSDYASGYHWGSCBCKAFKRSIOGH 180
 Db 181 ndylcpalnqctldknrrksqacrlrkcyegmwkcgsrrercgylvrrgrsadeqlh 240
 OY 181 NDYICPATNQTIDKRRKSCQACRLRKCYEGMWKCGSRRERCGYLRVRRORSADQLH 240
 Db 241 cagfakrsggghaprvrellldalspeqlvltlleaprhvllsrpsaiftasmmmsltk 300
 OY 241 CAGAKRSGGHAAPRVRELLDLALSPQLVLTLLLEAPRHVLLSRPSAIFTASMMMSLTk 300
 Db 301 ladelvmlswakkpfpvvelsfddqvrllscwmevllmgllmrsldhpgkllfapdl 360
 OY 301 LADKELVMIWAKKIPFVELSLFDQVRLLSCWMEVLMGMKMRSLDHGKLLFAPDL 360
 Db 361 vldfdegkveglleifdmllatstfrfrelklqkheylcvkamllnssmyplvtatqda 420
 OY 361 VLDRDEGKVEGLIEIFDMLLATTSFRRELKQHKCYLCVKAMILLNSSMYPLVATQDA 420
 Db 421 dsstklahllnavtdalvwiaksgisgqgmllanllmllshvrhshnkymehllmkk 480
 OY 421 DSSRLAHLNAVTDALVWIAKSGISSQOOSMRLANLMLLSHVRAHSNKGMEHLNMK 480
 Db 481 cknvpyvdllemnahvllrgckstfgscspaedsksgsqnpsq 530
 OY 481 CKNVYPVLDLLEMLNAHVLRGCKSTFGSCSPAEKSKSGSQNPSQ 530

RESULT 2
 ID W14724 standard; Protein; 485 AA.
 AC W14724:
 DT 08-JUN-1997 (first entry)
 DE Human oestrogen receptor beta (ER-beta).
 KW Ophan receptor; oestrogen receptor beta; ER-beta;
 KW nuclear receptor; prostate cancer; benign prostatic hyperplasia;
 KW ovary cancer; cardiovascular disease; osteoporosis;
 KW environment; pollutant.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT domain 104..169
 FT /label= DNA-binding_domain
 FT domain 260..457
 FT /label= Ligand-binding_domain
 PN W09709348-A2.
 PD 13-MAR-1997.
 PR 09-SEP-1996; E03933.
 PR 08-SEP-1995; GB-018272.
 PR 15-MAR-1996; GB-005550.
 PR 11-APR-1996; GB-007532.
 PR 08-MAY-1996; GB-009576.
 PA (KARO-) KARO BIO AB.
 PI Emma E, Gustafsson J, Kulper GG;
 DR N-PSDB: T62843.
 DR WPI: 97-192842/17.
 PT New isolated oestrogen receptor beta - used to develop prods. for
 PT treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
 PT disease
 PS Claim 1: Fig 13A; 45pp; English.
 CC A novel human orphan receptor (W14724) is related to the known
 CC oestrogen receptor ER-alpha, and has been designated ER-beta.
 CC It is an oestrogen receptor-related nuclear receptor. The ER-beta
 CC amino acid sequence was deduced from a cDNA clone (T62843) isolated
 CC from a human testis cDNA library. Rat, human and mouse ER-beta
 CC (W14723-25) can be used to isolate molecules for use in the
 CC treatment of cardiovascular diseases, central nervous system
 CC diseases, osteoporosis, prostate or ovarian cancer or benign
 CC prostatic hyperplasia and to test environmental chemicals for
 CC oestrogenic activity.
 SQ Sequence 485 AA;

Query Match 91.8%; Score 3582; DB 21; Length 485;
 Best Local Similarity 99.8%; Pred. No. 0.00e+00;
 Matches 484; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 mtfyspavmmslpsnvtlnlegpgrqtlspnvlwtpbphlspvlvhrqlshlyaeopks 60
 OY 46 MTFYSPAVMMSLPSNVTNLEGPGRQTSPNVLWTPBPHLSPLVHRQLSHLYAEOPKS 105
 Db 61 pwcarslehtlpvncetlkrkvsgnrcasvtpbpqskrdahfcavcsdyassghygvws 120
 OY 106 PWCARSLHTLPVNRRLTKRKVSGNRCASVTPBQSKRDHFCVAVCSGYHYGVWS 165
 Db 121 ceegkafkrszlgghndylcpatnqctldknrrksqacrlrkcyegmwkcgsrrercg 180
 OY 121 CEGKAFKRSIQGHNDYICPATNQTIDKRRKSCQACRLRKCYEGMWKCGSRRERCG 225
 Db 166 CEGKAFKRSIQGHNDYICPATNQTIDKRRKSCQACRLRKCYEGMWKCGSRRERCG 225
 OY 181 yrlvrrgrsadeqlhcgakrsgghaprvrellldalspeqlvltlleaprhvllsrp 240
 Db 226 YLRVRRORSADQLHCGAKRSGGHAAPRVRELLDLALSPQLVLTLLLEAPRHVLLSRP 285
 OY 241 sapfteaammslkladkelvmlswakkpfpvvelsfddqvrllscwmevllmgllmrs 300
 Db 286 SAPFTEASMMMSLTKLADKELVMIWAKKIPFVELSLFDQVRLLSCWMEVLMGMKMR 345
 OY 301 rsldhpgkllfapdlvldrdegkveglleifdmllatstfrfrelklqkheylcvkamll 360
 Db 346 RSIDHPGKLIFAPDLVDRDEGKVEGLIEIFDMLLATTSFRRELKQHKCYLCVKAMIL 405
 OY 361 lnssmyplvtatqdasstklahllnavtdalvwiaksgisgqgmllanllmllshv 420
 Db 406 LNSSMYPLVATQDASSRLAHLNAVTDALVWIAKSGISSQOOSMRLANLMLLSHV 465
 OY 421 rhasnkgmehllnmckknvpyvdllemnahvllrgckstfgscspaedsksgsq 480
 Db 466 RHASNKGMEHLNMCKKNVYPVLDLLEMLNAHVLRGCKSTFGSCSPAEKSKSGSQ 525
 OY 481 nlqsg 485
 Db 526 NPOSQ 530

RESULT 3
 ID W33212 standard; Protein; 477 AA.
 AC W33212:
 DT 20-APR-1998 (first entry)
 DE Human oestrogen receptor protein.
 KW Oestrogen receptor protein; steroid; alternative splicing; estradiol;
 KW estrone; estrinol; screening.
 OS Homo sapiens.
 FH Key
 FT domain 1-1997.
 FT EP-798378-A2.
 PD 01-OCT-1997.
 PR 25-MAR-1997; 200903.
 PR 22-NOV-1996; EP-203284.
 PR 26-MAR-1996; EP-200820.
 PA (ALKU) AKZO NOBEL NV.
 PI Dijkema R, Mosselman S;
 DR WPI: 97-473188/44.
 DR N-PSDB: T88412.
 PT DNA encoding estrogen receptor - useful in screening assay to
 PT identify novel ligands or hormonal analogues
 PS Claim 4; Page 18-20; 45pp; English.
 CC This sequence represents a novel oestrogen binding protein. The cDNA
 CC sequence which encodes this protein can be alternatively spliced
 CC resulting in the detection of additional transcripts (see T88413).
 CC This receptor is able to bind and be activated by estradiol, estrone and
 CC estrinol, can be used in a screening assay for the identification of new
 CC drugs e.g. novel ligands or hormonal analogues.
 SQ Sequence 477 AA;

Query Match 90.5%; Score 3532; DB 28; Length 477;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 mmyslpsnvtlnlegpgrqtlspnvlwtpbphlspvlvhrqlshlyaeopks 60
 OY 54 MMYSLPSNVTNLEGPGRQTSPNVLWTPBPHLSPLVHRQLSHLYAEOPKSEARSL 113

```

Db 61 ehlpvnrcltkrvsgnrcasvptpgskrdahfcavcasygyhyvwscecgkaff 120
      |||
Qy 114 EHLTPVNRCLTKRVSGNRCASVPTPGSKRDHAFCAVCASYGYHYVWSCGCKAFF 173
      |||
Db 121 krsigghndyicpatngctldknrrkscgacrlrkcyevmvgcgsrrercgyrlvrrgr 180
      |||
Qy 174 KRSTGCHNDYICPATNGCTLDKNRRKSCQACRLRKCEYGMVWGSGRRRCGRVLRVRQR 223
      |||
Db 181 sadeqlhcgakkrsgghaprvrellldalspeqlvlllaepphyllisrpaafteas 240
      |||
Qy 234 SADEQLHCAGAKRSGGHAAPRVRELLDLALSPQVLVTLLEAPPHVLSRPAAPTEAS 223
      |||
Db 241 mmmstlkladkelvnmisvaakkipgfvelslfdqvrllscwmevllmnglmwrsidhpgk 300
      |||
Qy 294 MMSLTKRLADKELVNMISVAAKKIPGFVELSLFDQVRLLSCWMEVLMNGLMWRSIDHPGK 353
      |||
Db 301 lfapdlvldrdcgkveglleifdmllatrsfrfrelklqhkelylcvkamllnsmyrl 360
      |||
Qy 354 LIFAPDLVLD RDGKCEGLLEIFDMLLATRSFRRELKLQHKELYLCVKAMLLNNSMYPL 413
      |||
Db 361 vlatqgadesrklahlinaavtdalvwwiaksgisgqgsurmlanllshvrhasnkgm 420
      |||
Qy 414 VYATQADSSRKLAHLNNAVTDALVWVIAKSGISSQOQSMRLANLMLLSHVHASNKGK 473
      |||
Db 421 ehllmckknvpyd11lemlnahvltgckssitgsecspadcskskegsqnpqsg 477
      |||
Qy 474 EHLNMCKKNVYPYD11LEMLNAHVLTGCKSSITGSECSPADCSKSKESQNPQSG 530
      |||

```

RESULT 4
ID W14723 standard; Protein: 485 AA.

```

AC W14723;
DE 08-JUN-1997 (first entry)
DE Rat oestrogen receptor beta (ER-beta).
KW Orphan receptor; oestrogen receptor beta; ER-beta;
KW nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW ovary cancer; cardiovascular disease; osteoporosis;
KW environment; pollutant.
OS Rattus sp.
FH Key Location/Qualifiers
FT domain 104..169
FT domain /label= DNA-binding_domain
FT domain 259..457
FT domain /label= Ligand-binding_domain
PN WO9709348-A2.
PD 13-MAR-1997.
PE 09-SEP-1996; E03933.
PR 08-SEP-1995; GB-018272.
PR 15-MAR-1996; GB-005550.
PR 11-APR-1996; GB-007532.
PR 08-MAY-1996; GB-009576.
PA (KARO-) KARO BIO AB.
PI Enmark E, Gustafsson J, Kuiper GG;
DR WPI: 97-192842/17.
DR N-PSDB: T62842.
PT New isolated oestrogen receptor beta - used to develop prods. for
PT treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT disease
PS Claim 1; Fig 1; 45pp; English.
CC A novel rat orphan receptor (W14723) is related to the known
CC oestrogen receptor ER-alpha, and has been designated ER-beta.
CC It is an oestrogen receptor-related nuclear receptor. The ER-beta
CC amino acid sequence was deduced from a cDNA clone (T62842) isolated
CC from a rat prostate cDNA library. Rat, human and mouse ER-beta
CC (W14723-25) can be used to isolate molecules for use in the
CC treatment of cardiovascular diseases, central nervous system
CC diseases, osteoporosis, prostate or ovarian cancer or benign
CC prostatic hyperplasia and to test environmental chemicals for
CC oestrogenic activity.
CC Sequence 485 AA:

```

Query Match 83.9%; Score 3274; DB 21; Length 485;
Best Local Similarity 88.7%; Pred. No. 0.00e+00;

```

Matches 430; Conservative 35; Mismatches 20; Indels 0; Gaps 0;
Db 1 mtfyspawmvywpgstslldggpvrlstspnlytpstglsplahqcslllyapokk 60
      |||
Qy 46 MTFYSPAWMVYWPGSTSLLDGGPVRLSTSPNLYTPSTGLSPLAHQCSLLIYAPOKS 105
      |||
Db 61 pwcearslehtlpvnrcltkrklsgscasvptspnkrdahfcpvcasygyhyvws 120
      |||
Qy 106 PWCERSLEHTLPVNRCLTKRKYSGRNRCASVPTGPGSKRDHAFCAVCASYGYHYVWS 165
      |||
Db 121 cegckaftrsisghndyicpatngctldknrrkscgacrlrkcyevmvgcgsrrercg 180
      |||
Qy 166 CEGCKAFTRFSIGHNDYICPATNGCTLDKNRRKSCQACRLRKCEYGMVWGKSGRRRCG 225
      |||
Db 181 yrlvrrorsadepohcagakrakgghaprvrellldalspeqlvlllaepphyllisr 240
      |||
Qy 226 YRLVRRORSADepohCAGAKRKGGHAPRVRELLDLALSPQVLVTLLEAPPHVLSR 285
      |||
Db 241 smpfteasmmmsltkladkelvnmisvaakkipgfvelslfdqvrllscwmevllmnglmw 300
      |||
Qy 286 SAMPFTEASMMMSLTKLADKELVNMISVAAKKIPGFVELSLFDQVRLLSCWMEVLMNGLMW 345
      |||
Db 301 rsidhpgkllfapdlvldrdcgkveglleifdmllatrsfrfrelklqhkelylcvkamll 360
      |||
Qy 346 RSIDHPGKLLFAPDLVLD RDGKCEGLLEIFDMLLATRSFRRELKLQHKELYLCVKAMIT 405
      |||
Db 361 lnsmyrlasangeaessrklhlinaavtdalvwwiaksgisgqgsurmlanllshv 420
      |||
Qy 406 LNSMYRPLVATQADSSRKLAHLNNAVTDALVWVIAKSGISSQOQSMRLANLMLLSHV 465
      |||
Db 421 rthsnkgmehllsmckknvpyd11lemlnahvltgckssitgsecspadcskskegsq 480
      |||
Qy 466 RTHSNKGMEHLLSMCKKNVYPYD11LEMLNAHVLTGCKSSITGSECSPADCSKSKESQ 525
      |||
Db 481 nlqsg 485
      |||
Qy 526 NPQSG 530
      |||

```

RESULT 5
ID W14725 standard; Protein: 485 AA.

```

AC W14725;
DE 08-JUN-1997 (first entry)
DE Mouse oestrogen receptor beta (ER-beta).
KW Orphan receptor; oestrogen receptor beta; ER-beta;
KW nuclear receptor; prostate cancer; benign prostatic hyperplasia;
KW ovary cancer; cardiovascular disease; osteoporosis;
KW environment; pollutant.
OS Mus sp.
FH Key Location/Qualifiers
FT domain 104..169
FT domain /label= DNA-binding_domain
FT domain 260..457
FT domain /label= Ligand-binding_domain
PN WO9709348-A2.
PD 13-MAR-1997.
PE 09-SEP-1996; E03933.
PR 08-SEP-1995; GB-018272.
PR 15-MAR-1996; GB-005550.
PR 11-APR-1996; GB-007532.
PR 08-MAY-1996; GB-009576.
PA (KARO-) KARO BIO AB.
PI Enmark E, Gustafsson J, Kuiper GG;
DR WPI: 97-192842/17.
DR N-PSDB: T62844.
PT New isolated oestrogen receptor beta - used to develop prods. for
PT treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
PT disease
PS Claim 1; Fig 14a; 45pp; English.
CC A novel mouse orphan receptor (W14725) is related to the known
CC oestrogen receptor ER-alpha, and has been designated ER-beta.
CC It is an oestrogen receptor-related nuclear receptor. The ER-beta
CC amino acid sequence was deduced from a cDNA clone (T62844).

```

CC Rat, human and mouse ER-beta (M44723-25) can be used to isolate
CC molecules for use in the treatment of cardiovascular diseases,
CC central nervous system diseases, osteoporosis, prostate or ovarian
CC cancer or benign prostatic hyperplasia and to test environmental
CC chemicals for oestrogenic activity.
SQ Sequence 485 AA;

Query Match

[illegible]

RESULT	6	
ID	W33213	standard; Protein; 416 AA.
AC	W33213.	
DT	20-APR-1998	(first entry)
DE	Human oestrogen receptor protein exon 8 splice variant.	
KW	Oestrogen receptor protein; steroid; alternative splicing; estradiol; estrone; estriol; screening.	
OS	Homo sapiens.	
PN	EP-798378-A2.	
PD	01-OCT-1997.	
PF	25-MAR-1997.	200903.
PR	22-NOV-1996.	EP-203284.
PR	26-MAR-1996.	EP-200820.
PA	(ALKU) AKZO NOBEL NV.	
PI	Dijkema R, Mosselman S;	
DR	WPI; 97-473188/44.	
DR	N-PDSB; T88413.	
PT	DNA encoding estrogen receptor - useful in screening assay to	
PT	Identify novel ligands or hormonal analogues	
PS	Claim 4.; Page 20-22; 45bp; English.	
CC	This partial sequence represents a splice variant of a novel oestrogen	
CC	binding protein, detected by screening a human testis cDNA library.	
CC	This protein contains an alternative exon 8 (exon 8B) of the novel	
CC	oestrogen receptor represented. As a consequence of the introduction of	
CC	this exon through an alternative splicing reaction, the reading frame	

CC encoding the novel receptor is immediately terminated, creating a
CC truncation of the carboxy terminus of the novel receptor. This receptor
CC is able to bind and be activated by estradiol, estrone and estril, can be
CC used in a screening assay for the identification of new drugs e.g. novel
CC ligands or hormonal analogues. This variant does not contain an Af-2
CC region and therefore probably lacks the ability to modulate transcription
CC of target genes in a ligand dependent fashion.
Sequence 416 AA;

Query Match	79.48;	Score 3099;	DB 27,	Length 416;
Best Local Similarity	100.0%;	Pred. No. 4.32e-293;		
Matches 415; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Db	1	myisipnvnthlegprrcttsprvltwptghslplvthqjshlaepokspwcearsl	60
Oy	54	MYISIPSNVNLLEGPPRRCTTSPRVLTWPTGHSPLVTHQJSHLAEPOKSPWCEARSL	113
Db	61	ehltipvnetlkrvsgnrcaspytgpsskrdahtfcavcsdyasgylyuawsccegckaff	120
Oy	114	EHLTPVNETLKRVSNGNRCASPYTGPSKRDHAFCAVCSDYASGYLYUAWSCGCKAFF	173
Db	121	krsiqghndyicpeltngctldknrtksccgacrltkcyevmvmvkcgsrtercgyrlvtrqr	180
Oy	174	KRSIQGHNDYICPATNCTTIDKNRKRKSCQACRLKCYEVMVMVKGSRTERCGYRLVTRQR	233
Db	181	sadeglhaoagkrakssggghaprvrrllldalspeglvllleaepphvlisrpsafteas	240
Oy	234	SADegLHOGAKRKSggGhAPRVRRLLDLASpEGLVLLLEaEPPhVLISrPSaPTEAS	293
Db	241	mmmsjtkladelvhnmsiwaaklpgfvelslfdqvrlllescmewlmmjlmwrsldhpgk	300
Oy	294	MMMSITKLADKELVHNMISMAKRGFVELSLFDQVRLLLESCMEWVLMMLMWSIDHPGK	353
Db	301	llfapdlvldtdegkcveglleifdmllatstrfrelkldqkeylcvkamillosmsypl	360
Oy	354	LLFAPDLVLDDEBKCVEGLEIFEDMLLATSTRFRELKLDHKEYLCKAMILLNMSYPL	413
Db	361	vtatogdadsrklhlllnavrdalvwrlaksgjassggsgsmrllnlllshvtrha	415
Oy	414	VTATODADSSRKLHLNLNAVTDALVWRLAKSGJSSGQOOSMRLLNMLLSHVTRHA	468

RESULT 7
ID W33214 strandcd: Protein; 418 AA.
AC W33214;
DT 20-APR-1998 (first entry)
DE Human oestrogen receptor protein splice variant (exon 8c).
KW Oestrogen receptor protein; steroid; alternative splicing; estradiol;
KW estrone; estril; screening.
OS Homo sapiens.
PN Ep-798378-A2.
PD 01-OCT-1997.
PF 25-MAR-1997; 200903.
PR 22-NOV-1996; EP-203284.
PR 26-MAR-1996; EP-200820.
PA (ALKU) AKZO NOBEL NV.
PI Dijkema R, Mosselman S;
DR WP1; 97-473188/44.
DR N-PSDB; T88414.
PT DNA encoding estrogen receptor - useful in screening assay to
PT identify novel ligands or hormonal analogues
PS Claim 4; Page 30-35; 45pp; English.
CC This sequence represents a splice variant of a novel oestrogen binding
CC protein isolated from human thymus tissue. This protein contains an
CC alternative exon 8 (exon 8C) of the novel estrogen receptor represented
CC in T88412. This novel receptor is able to bind and be activated by
CC estradiol, estone and estril, can be used in a screening assay for the
CC identification of new drugs e.g. novel ligands or hormonal analogues.
CC This variant does not contain an AF-2 region and therefore probably
CC lacks the ability to modulate transcription of target genes in a ligand
CC dependent fashion.
SQ Sequence 418 AA;

Query Match 79.4%; Score 3099; DB 28; Length 418;
 Best Local Similarity 100.0%; Pred. No. 4.32e-293;
 Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mnyispsnvnleegpqrptspnvlwptpghlspvlvhrhgllvyeepkspcearsl 60
 |||||
 QY 54 MNYSIPSNVNTLEGGPRQTTSPVNLWPTPEHLSPLVYVHRLSHLVEPKSPMCEARSL 113
 |||||
 Db 61 ehtlpvnetlkrkvsngarcsapvtgprskrdahfcavcsgdyaghygvscgckaff 120
 |||||
 QY 114 EHTLPVNETLKRKVSNGARCSAPVTGPRSKRDHFAVCASDYAGHYGVSCGCKAFF 173
 |||||
 Db 121 krsiqgmndyicpatngctldknrrkscgacrlkcyevgmwkcgsrrecgylvrqr 180
 |||||
 QY 174 KRSIQGMNDYICPATNGCTLDKNRRKSCGACRLKCYEVGMWKCGRRECGRYLVRRQR 233
 |||||
 Db 181 sadeqlhacgkkrsgggharvrrelldalspeqlvltlleaephylisrpsapfcea 240
 |||||
 QY 234 SADEQLHACGKAKRSGGHARVRRELLDALSPEDLVLTLEAEPHYLISRPSAPFEAS 293
 |||||
 Db 241 mmsstklackelvhmismwakkipgfvelsifdgvrllscemewlmvglmwrsidhpgk 300
 |||||
 QY 294 MMSSTKLADKELVHMTSMWAKKIPGFVELSIFDQVRLLESCMMEVLMGLMKMRSIDHPRK 353
 |||||
 Db 301 lifapdlvldrdegkveglleifdmllatsrfrelkqhkeylcvkamllnssmypy 360
 |||||
 QY 354 LIFAPDLVLRDEGKVEGLEIFDMLLATTSRRELKQKEYLCVKAMILLNSSMYP 413
 |||||
 Db 361 vtttqddssrklahlnavtdaivwriaksglsgsggmrlanllmshvrha 415
 |||||
 QY 414 VTATQDDSSRKLALHNLNAVTDALVWRIASGLSSGSGMRLANLLMLSHVRA 468
 |||||

RESULT 8

W33211, standard; protein; 233 AA.
 AC W33211;
 DT 20-APR-1998 (first entry)
 DE Human oestrogen receptor protein ligand binding domain.
 KW Oestrogen receptor protein; steroid; estradiol; estrone; estrinol;
 KM screening; ligand binding domain.
 OS Homo sapiens.
 PN EP-798378-A2.
 PD 01-OCT-1997.
 PE 25-MAR-1997; 200903.
 PR 22-NOV-1996; EP-203284.
 PS 26-MAR-1996; EP-200820.
 PA (ALKT) AKZO NOBEL NV.
 PI Dijkema R, Mosselman S;
 WP1; 97-473187/44.
 DR N-PSDB; T88412.
 PT DNA encoding estrogen receptor - useful in screening assay to
 identify novel ligands or hormonal analogues
 PS Claim 4; Page 16-17; 45pp; English.
 CC This sequence encodes the ligand binding domain of a novel
 CC oestrogen binding protein, detected by screening a human testis
 CC cDNA library. This receptor is able to bind and be activated by
 CC estradiol, estrone and estrinol, can be used in a screening assay for
 CC the identification of new drugs e.g. novel ligands or hormonal
 CC analogues.
 CC Sequence 233 AA;

Query Match 43.9%; Score 1714; DB 27; Length 233;
 Best Local Similarity 99.6%; Pred. No. 2.24e-154;
 Matches 232; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 lvltlleaephylisrpsapfceaasmmsltklackelvhmismwakkipgfvelsifdg 60
 |||||
 QY 268 LVLTLEAEPHYLISRPSAPFEASMMSLTKLADKELVHMTSMWAKKIPGFVELSIFDQ 327
 |||||
 Db 61 vrllescmewlmvglmwrsidhpgklifapdlvldrdegkveglleifdmllatsrf 120
 |||||
 QY 328 VRLLESCMEVLMGLMKMRSIDHPRKLIAPDLVLRDEGKVEGLEIFDMLLATTSRF 387
 |||||

Db 121 relkqhkeylcvkamllnssmypyvtatcgadssrklahlnavtdaivwriaksgis 180
 |||||
 QY 388 RELKQKEYLCVKAMILLNSSMYPVTATQDSSRKLALHNLNAVTDALVWRIAKSGIS 447
 |||||

Db 181 sqggsmtlanllmshvrhaakgmehllnmcknvpywdlllemnahvl 233
 |||||
 QY 448 SQOOSMRLANLLMLSHVRAHAKGMHLLNMCKNVVYDLEMLLMAHVL 500
 |||||

RESULT 9

P70543, standard; protein; 595 AA.
 AC P70543;
 DT 24-FEB-1991 (first entry)
 DE Human oestrogen receptor protein.
 KW Human oestrogen receptor protein; expression system; assay.
 OS Homo sapiens.
 PN W08705049-A.
 PD 27-AUG-1987.
 PE 18-FEB-1987; U00341.
 PR 20-FEB-1986; US-833829.
 PA (CALB-) California Biotechnology Inc.
 PI Shine J;
 WP1; 87-250213/35.
 DR N-PSDB; N70880.
 PT Expression system for vertebrate steroid receptor protein - comprising
 PT DNA sequence encoding the protein linked to control sequences in
 PT eucaryotic hosts.
 PS Disclosure; Fig. 1-1 - 1-2; 26pp; English.
 CC Steroid receptor protein;
 CC The protein is expressed in an expression system from cDNA operably
 CC linked to control sequences compatible with eukaryotic host cells.
 CC This method allows expression under conditions which favour appropriate
 CC post-translational processing. It also produces large amts. of purified
 CC protein useful in the design of agonist and antagonist cpsds. for study
 CC of the mechanism of action of the steroid binding proteins in general,
 CC and for use in diagnostic assays for the proteins or antibodies to
 CC them. These assays are important in, eg the diagnosis of tumour
 CC sensitivities to steroid metabolism. Suitable host cells are VERO,
 CC HeLa and CHO cells.
 CC Sequence 595 AA;

Query Match 42.8%; Score 1671; DB 2; Length 595;
 Best Local Similarity 56.3%; Pred. No. 4.30e-150;
 Matches 227; Conservative 98; Mismatches 61; Indels 17; Gaps 13;

Db 178 saketrycavcndyagsgyhygvscgckafkrsiqgmndyicpatngctldknrrksc 237
 |||||
 QY 142 SKRDHRCVACSDYASGYHYGVSCGCKAFKRSIQGMNDYICPATNGCTLDKNRRKSC 201
 |||||
 Db 238 gaerlkyevgmwkmkgikldrrgtrmlkhrqrddeggrgvegsagdmraanlwpsslm 297
 |||||
 QY 202 QACRLRKYCEVGMWKCGRRECGRYLVRR--RORSAD-QLHRCAGRA-KRSGG--HAP-R 254
 |||||
 Db 298 ikxskksnslstldagmwsalldaapp-llseybptprfseasmglltnladrelvh 356
 |||||
 QY 255 V-KELL-LD-ALSPEDLVLTLEAEPHYLISR--PSAPFEASMMSLTKLADKELVH 308
 |||||
 Db 357 mlnwkrvpgfvdltldghyllleacawlellmiglvrsmehpvkllfapanllidtnqgk 416
 |||||
 QY 309 MISWAKKIPGFVELSIFDQVRLLESCMMEVLMGLMKMRSIDHPRKLIAPDLVLRDEGK 368
 |||||
 Db 417 cvegmveifdmllatsrfrmmnlgeefvclskslilnsgvytlfslstlsleekdhln 476
 |||||
 QY 369 CVEGIEIIFDMLLATTSRRELKQKEYLCVKAMILLNSSMYPVTATQDA-DSSRKL 427
 |||||
 Db 477 rvidktldtlhmakarltlqgghqrlqlllshtrmsnkgmehlymcknvpywdlllemnahvl 536
 |||||
 QY 428 HLNAVTDALVWRIAKSGISQOOSMRLANLLMLSHVRAHAKGMHLLNMCKNVVY 487
 |||||
 Db 537 ydlillemlahrlha-ptsrgasveetdshlatatsashs 578
 |||||
 QY 488 YDLLEMLNHLVLRCKSSITGSECPADDSK-SKRGSONPOS 529
 |||||

```

RESULT 10
ID P80925 standard; protein; 334 AA.
AC P80925;
DE 29-DEC-1990 (first entry)
DE Sequence of the human oestrogen receptor
KW Hormone receptor; hormone-binding.
OS Homo sapiens.
PN WO8803168-A.
PD 05-MAY-1988.
PE 23-OCT-1987; US-108471.
PR (Salk) Salk Inst for Biol Stud.
PI Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR WPI; 88-133242/19.
DR N-PSDB; N80922.
PT Recombinant DNA encoding hormone receptors -
PT comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT and novel hormone receptors
PS Example; Fig III-2; 243pp; English.
CC DNAs encoding hormone receptors and the hormone receptors themselves are
CC claimed. The DNA can be used to make the hormone receptor proteins and
CC functional modified forms in quantities not previously possible. The
CC receptor proteins can be used to screen cpts. for receptor-agonist or
CC receptor-antagonist activity. They can also be used in diagnostic assays.
SQ Sequence 334 AA;

Query Match 36.5%; Score 1424; DB 2; Length 334;
Best Local Similarity 57.4%; Pred. No. 1,58e-125;
Matches 193; Conservative 80; Mismatches 47; Indels 16; Gaps 12;

Db 1 ycaavcdasygyhygvscgcgcafkftrs-qghndymcpatnctidknrtkscgacrlr 59
QY :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
148 FCACVSDYASGYHYGVSCGCGCAFFKRSIQGHNNDYICPATNOCTIDKNRKSQCACRLR 207
Db 60 kcytgmkmkgkikrttgrgmikhrqddggrgvgvsagdmzamlwpsrlmkikrsk 119
QY |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
208 KCYEGVMKCGSRRECGRLVR--RQNSADQ-LHCAGKA-KRSGG--HAF-RV-RELL 259
Db 120 nsalstladgmvsalldaerp-ilyeydprtrfseasmngllnsladrelvhlmwak 178
QY |:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
260 --LD-ALSPEDQVLTLLAEFPHVLSR--PSAPFTEASMMSLTKRLADKELVHISMNAK 314
Db 179 tvpgrfdvltldqvhlllecaulelmiglvtrsmehpvyklfapnllldrngkcvsgmv 238
QY ::|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
315 KIPGEVELSLFDQVRLSLSCWMEVLMGMLMRSIDHPKLIAPDLVDRDGGKCEGTL 374
Db 239 eifdmllstsrfrmmnlqgeefvclkslllinsgyvflstlksleekdhivldki 298
QY |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
375 EIEFDMLATTSRERELKLOHKEYLCVKAMITLNLSSMYPLVATODA-DSSRKLHLNNAV 433
Db 299 tdtllhmakagltlqqghqrlaqllllshirfms 334
QY |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
434 TDAIWMVIAKSGSISSQOOSMRLANLMLLSHVRHNS 469

RESULT 11
ID P80931 standard; protein; 433 AA.
AC P80931;
DE 29-DEC-1990 (first entry)
DE Sequence of human estrogen-related receptor protein (hERR2)
KW Hormone receptor; hormone-binding; transcription activation.
OS Homo sapiens.
PN WO8803168-A.
PD 05-MAY-1988.
PE 23-OCT-1987; US-108471.
PR (SALK) Salk Inst for Biol Stud.
PI Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR WPI; 88-133242/19.
DR N-PSDB; N80922.
PT Recombinant DNA encoding hormone receptors -
PT comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT and novel hormone receptors
PS Claim 17; Fig V-2(B)-1 and -2; 243pp; English.
CC DNAs encoding hormone receptors and the hormone receptors themselves are
CC claimed. The DNA can be used to make the hormone receptor proteins and
CC functional modified forms in quantities not previously possible. The
CC receptor proteins can be used to screen cpts. for receptor-agonist or
CC receptor-antagonist activity. They can also be used in diagnostic assays.
CC Also claimed is pure DNA encoding estrogen-related receptors hERR1 and
CC hERR2 and hERR1 and hERR2. The new hERR1 and hERR2 receptors will provide
CC the basis for development of an assay system that will lead to the
CC identification of novel hormones.
SQ Sequence 433 AA;

Query Match 20.7%; Score 807; DB 2; Length 433;
Best Local Similarity 39.5%; Pred. No. 1,04e-64;
Matches 140; Conservative 84; Mismatches 108; Indels 22; Gaps 14;

Db 101 rlcvcddasgyhygvscgcgcafkftrtlignlieyscpatneceltkrtkscgacrf 160
QY ::|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
147 HFCACVSDYASGYHYGVSCGCGCAFFKRSIQGHNNDYICPATNOCTIDKNRKSQCACRL 206
Db 161 mcklvgmlkegvldtr--r-ggrqky-krldsenspylsqispakpklkivs-- 214
QY |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
207 RKCYEGVMKCGSRRECGRLVRGRSADQLHCAGKARSGGHAPRVRELLDALSPE 266
Db 215 yl-l-vaepdklyam--ppddvpegdialtlcladrelvfliswakhlpfnsnlgd 271
QY |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
267 QLVLTLLAEFPHVLSRPSAPFTEASMMSLTKRLADKELVHISMNAKPIGFEVLSLFD 326
Db 272 gmslllgasmealllgivysrlypydklayaedymdeesrlv-glllelyralqlvtr 330
QY |:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
327 QVRLESQWMEVLMGMLMRSIDHPGKLIAPDLVDRDGGKCEGTLFDMLATTSR 386
Db 331 ykklkveefvmlkalalansdsmy--l--enleavqkldllh--ealg-dyels 380
QY ::|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
387 FRELKIDHEKVELCYKAMILLNS-SMYPVLTATQDADSSRLAHNLNVLTALVWIAKSG 445
Db 381 -qrheeptragklllclprrlqrtaakavqfyvklgkvpkmhlflemleakv 433
QY ::|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
446 ISSQOOSMRLANLMLLSHVRHNSKGMHEHLNKKCRNVVPVYDLEMLNNAV 499

RESULT 12
ID P80930 standard; protein; 521 AA.
AC P80930;
DE 10-MAR-1993 (revised)
DE 29-DEC-1990 (first entry)
DE Sequence of human estrogen-related receptor protein (hERR1)
KW Hormone receptor; hormone-binding; transcription activation.
OS Homo sapiens.
PN WO8803168-A.
PD 05-MAY-1988.
PE 23-OCT-1987; US-108471.
PR (SALK) Salk Inst for Biol Stud.
PI Evans RM, Weinberger CA, Hollenberg SM, Giguere V;
DR WPI; 88-133242/19.
DR N-PSDB; N80921.
PT Recombinant DNA encoding hormone receptors -
PT comprising glucocorticoid, mineralocorticoid, thyroid hormone
PT and novel hormone receptors
PS Claim 17; Fig V-1(B)-1 and -2; 243pp; English.
CC DNAs encoding hormone receptors and the hormone receptors themselves are
CC claimed. The DNA can be used to make the hormone receptor proteins and
CC functional modified forms in quantities not previously possible. The
CC receptor proteins can be used to screen cpts. for receptor-agonist or
CC receptor-antagonist activity. They can also be used in diagnostic assays.
CC Also claimed is pure DNA encoding estrogen-related receptors hERR1 and
CC hERR2 and hERR1 and hERR2. The new hERR1 and hERR2 receptors will provide
CC the basis for development of an assay system that will lead to the
CC identification of novel hormones.
SQ Sequence 521 AA;

Query Match 20.5%; Score 800; DB 2; Length 521;

```